Investigation into the Interactions between thermal management, lubrication and control systems of a diesel engine

Burke, Richard

Award date:
2011

Awarding institution:
University of Bath

Investigation into the Interactions between Thermal Management, Lubrication and Control Systems of a Diesel engine

Richard David Burke

A thesis submitted for the degree of Doctor of Philosophy
University of Bath
Department of Mechanical Engineering
September 2011

Abstract

Engine thermal and lubricant systems have only recently been a serious focus in engine design and in general remain under passive control. The introduction of active control has shown benefits in fuel consumption during the engine warm-up period, however there is a lack of rigorous calibration of these devices in conjunction with other engine systems.

For these systems, benefits in fuel consumption (FC) are small and accurate measurement systems are required. Analysis of both FC and NO_x emissions measurements processes was conducted and showed typical errors of 1% in FC from thermal expansion and 2% in NO_x per g/kg change in absolute humidity. Correction factors were derived both empirically and from first principles to account for these disturbances. These improvements are applicable to the majority of experimental facilities and will be essential as future engine developments are expected to be achieved through small incremental steps.

Using prototype hardware installed on a production 2.4L Diesel engine, methodologies for optimising the design, control and integration of these systems were demonstrated. Design of experiments (DoE) based approaches were used to model the engine behaviour under transient conditions. A subsequent optimisation procedure demonstrated a 3.2% reduction in FC during warm-up from 25^{o}C under iso-NO_x conditions. This complemented a 4% reduction from reduced oil pumping work using a variable displacement pump.

A combination of classical DoE and transient testing allowed the dynamic behaviour of the engine to be captured empirically when prototype hardware is available. Furthermore, the enhancement of dynamic DoE approaches to include the thermal condition of the engine can produce models that, when combined with other available simulation packages, offer a tool for design optimisation when hardware is not available. These modelling approaches are applicable to a wide number of problems to evaluate design considerations at different stages of the engine development process. These allow the transient thermal behaviour of the engine to be captured, significantly enhancing conventional model based calibration approaches.

Acknowledgements

I would like to express my gratitude to Dr Chris Brace and Prof Gary Hawley for their supervision, guidance, approachability and support throughout this work. I would also like to thank Dr Sam Akehurst for his dedicated assistance in test cell operation – '*he is braver than I am when in control of an engine*'.

I would also like to thank Andy Lewis, Allan Cox and Sam Hurley for their efforts in designing, installing, operating and maintaining a well performing, functional test facility, regardless the many failures I have caused.

I must thank colleagues at the Ford Motor Company, Ian Pegg and Roland Stark, and at the University of Nottingham, Prof Paul Shayler and Jean-Paul Zammit, for their specialised technical advice. The experimental aspect of this work was conducted with the financial support of the Technology Strategy Board, Ford Motor Company, BP and Mahle Powertrain and their funding is acknowledged.

I am grateful to Dr Wolf Baumann of IAV GmbH who has spent considerable time teaching me the theoretical and practical aspects of dynamic design of experiments, as well as supplying the software tools necessary for this work. I also thank Ralf Wascheck for my inclusion in this aspect of the project.

I would also like to thank all my other colleagues at the PVRC for their support and advice throughout this project, Dr Chris Bannister, Dr Kevin Robinson, Ed, Joe, Mitch, Apiwat, Chris, Bjoern, Adam, Kai and Hasan.

It has been a pleasure working with you all.

Finally I would like to thank my good friends Pete and Phil for always being there. I am grateful for the support of my whole family, in particular Auntie Linda who has carefully read through large sections of this document. My darling Keeley, who has stood by me throughout and provided daily support, motivation and statistical tuition without which I would be stuck at *'Statistics-chapter 1'*. Last but not least I will thank my mum, dad and sister, Julie, Dave and Joanne for their unconditional support and for giving me the best possible opportunities in life to get this far.

Thank you all so much!

Table of contents

List of figures

List of tables

Notation

Abbreviations

Adj R^2	Adjusted R^2
APRBS	Amplitude modulated pseudo random binary signal
ATDC	After top dead centre
ATM	Active thermal management
BDC	Bottom dead centre
BMEP	Brake mean effective pressure
Bsfc	Brake specific fuel consumption
BTDC	Before top dead centre
CA	Crank Angle
CB	Carbon balance
CFD	Computational fluid dynamics
CI	Confidence interval
CO	Carbon monoxide
CO_2	Carbon Dioxide
DoE	Design of Experiments
DPF	Diesel particulate filter
ECU	Engine control unit
EGR	Exhaust gas recirculation
EOI	End of injection
EUDC	Extra-urban drive cycle
FC	Fuel consumption
FM	Friction modifiers
FMEP	Friction mean effective pressure
Grav	Gravimetric
H/X	Heat exchanger
HC	Hydrocarbons
HX	Heat exchanger
IC	Internal combustion
IMEP	Indicated mean effective pressure
N_2O	Nitrous oxide
NEDC	New European Drive Cycle
NO	Nitric oxide
NO_x	Oxides of Nitrogen
OLS	Orthogonal least squares
PAH	Polycyclic aromatic hydrocarbon
PM	Particulate Matter
PRESS	Predicted residual error sum of squares
PRT	Pressure regulated thermostat
R^2	Coefficient of determination
RBF	Radial base function
REA	Rotary electric actuator
RMSE	Residual mean square error
ROHR	Rate of heat release
SFC	Specific fuel consumption
SOI	Start of injection
TDC	Top dead centre
Temp	Temperature
TMS	Thermal management system
UDC	Urban drive cycle
VFOP	Variable flow oil pump
VGT	Variable geometry turbocharger

Mathematical symbols and subscripts

Statistical modelling (chapters 3 and 9)

Symbols

k	Time step
n	Number of test points
p	Probability
s	Standard deviation
$t_{n,95\%}$	Student's t value for n degrees of freedom and p=0.95
x	Independent input variable
y	Measured output
$\hat{y}$	Predicted output
$\overline{y}$	Mean output

Subscripts

1…n	independent input number

FC and emissions correction factors (chapter 3)

Symbols

FC	Fuel consumption	g
H_a	Absolute ambient humidity	$g/kg_{dry\ air}$
T_a	Ambient temperature	K
c_x	Emissions concentration of species X	% or ppm
f_{fw}	Volume increase factor from combustion	
k	Correction factor	
m	Mass	g
$\dot{m}$	Mass flow	kg/s
v	Volume	m^3
$\dot{v}$	Volume flow rate	m^3/s
w	Weight ratio	%
$\rho_{RATIO,X}$	Relative density of emissions species X	

Subscripts

C	Carbon
CB	Carbon balance
CO	Carbon monoxide
CO_2	Carbon dioxide
N	Nitrogen
NO_x	Oxides of nitrogen
O	Oxygen
bea	Beaker
cir	Fuel supply circuit
$cond$	Emissions analyser condenser
$corr$	Corrected
ex	Exhaust flow
dry	Without water vapour
end	end of test period
f	fuel
mes	measured
$start$	start of test period
t	time
wet	Including water vapour

Measurement systems (chapter 4)

Symbols

P	Power	W
Q	Heat	J
R		
T	Temperature, Torque	K or ^oC, Nm
U	Internal energy	J
V	Volume	m^3
V_s	Swept volume	m^3
W	Work	J
a, b	Regression constants	
c_p	Specific heat capacity at constant pressure	
c_v	Specific heat capacity at constant volume	
h	Specific enthalpy	J/kg
h_c	Convective heat transfer coefficient	W/m^2k
k	Thermal conductivity	W/mk
m	Mass	kg
$\dot{m}$	Mass flow	kg/s
p	pressure	bar or Pa

q	Heat flux	W/m^2
r	Radius	m
ω	Engine speed	rad/s

Subscripts

b	Bore
bb	Blow-by
c	Coolant
combustion	Combustion heat release
cy	Cylinder
ex	Exhaust
f	Fuel
in	Intake
Wall	Heat loss
wc	coolant wall interface

Further Publications

The work that will be presented has led to the publication or a number of articles in peer reviewed journals, at internationally recognised conferences and presented at technical seminars. A list of these publications is provided below and references to these will be provided as necessary within the chapters of this thesis.

Peer reviewed journal articles:

Critical Analysis of on-engine fuel consumption measurement

R.D.Burke et al., Proc. IMechE Vol. 225, Part D: J. Auto. Eng. (2011), issue 6, pp.829-844.

Systems Approach to the improvement of Engine Warm-up Behaviour

R.D.Burke et al., Proc. IMechE Vol. 225, Part D: J. Auto. Eng. (2011), issue 2, pp.190-205.

Review of systems analysis of the interactions of thermal, lubricant and combustion processes of Diesel engines

R.D.Burke et al., Proc. IMechE Vol. 224, Part D: J. Auto. Eng. (2010), issue 5, pp.681-704.

Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing

C.J. Brace, R.D.Burke et al., Proc. IMechE Vol. 223, Part D: J. Auto. Eng. (2009), issue 9, pp.1163-1177.

Refereed conference articles

Analysis of energy flows in engine coolant structure and lubricant during warm-up

R.D. Burke et al., presented at Vehicle Thermal Management Systems (VTMS) 2011 (Gaydon, UK, 16-19 May 2011)

The Effects of engine thermal management on performance, emissions and fuel consumption

R.D.Burke et al., presented at the SAE world congress 2010 (Detroit, US, April 2010), SAE paper number 2010-01-0802

Non-refereed conference articles

Practical approach to thermo-dynamic modelling of Diesel engine emissions

R.D. Burke et al., Presented at the 6th conference on Design of experiments in engine development (Berlin, Germany, 23-24 May 2011)

Engine and combustion stability over extended operating range using engine thermal management procedure

R.D.Burke et al., presented at the SIA Diesel Engine international conference (Rouen, France May 2010)

A Design of Experiments approach to the control of chassis dynamometer testing error

C.J. Brace, R.D Burke et al., presenter at the 5th conference on Design of Experiments in engine development (Berlin, Germany, June 2009)

Technical Seminar

Complete Drive cycle In-cylinder pressure measurement

R.D. Burke et al., ETAS Driving innovation 2010 (Donnington, UK, September 2010)

Chapter 1 - *Introduction*

Diesel engine cooling and lubricant systems will contribute to reducing emissions and fuel consumption through enhanced design functionalities. These new concepts will require active control from the engine control computer and the software will need to be enhanced to include these features. This work will establish a model-based calibration procedure for these systems, targeting the engine behaviour from cold start.

This chapter aims to lay out the background for the research presented in this thesis. A view of the current market place for automotive powertrain development will be presented from a commercial, environmental and legislative perspective. Engine parasitic losses will then be defined with a focus on the main theme of this thesis.

Based on this background, the aim and principle objectives will be presented and the areas of contribution and impact. Finally the layout of this thesis will be detailed.

1. *Background and motivation*

1.1. Driving factors

In most regions of the globe, regulations on vehicle tailpipe emissions have been in place for a number of years [1, 2]. In general the regulated emissions species are carbon monoxide (CO), unburned hydrocarbons (HC), oxides of nitrogen (NO_x) and smoke or particulate matter (PM). The limits are generally weight dependent to account for different vehicle uses and in the case of heavy duty trucks are power specific. These emissions trends have been met over the past years through various engine and after-treatment system improvements, however over this time carbon dioxide (CO_2) emissions have tended to increase, or at best stagnate[1]. CO_2 is a natural product of the combustion of carbon fuels and CO_2 reduction goes hand in hand with reduced fuel consumption.

Faced with the problems of the suspected impact of carbon dioxide on global warming, increasing fuel prices and depletion of reserves, authorities have recently moved towards the regulation of CO_2 emissions to encourage manufacturers to increase the efficiency of their products [3, 4]. In the European Union this will be a limit on CO_2 emissions[2] whereas proposals in the US and Japan relate to weight averaged fuel consumption [5]. With these upcoming regulations on fuel consumption and CO_2, in addition to continued tightening of other emissions regulations, it is widely accepted that meeting these targets is the biggest challenge facing the automotive industry.

1.2. Route to a low carbon future

Faced with the challenge above, it is widely accepted that the powertrain will play a major role in this development, and all manufacturers have laid out their development "roadmaps" for future vehicle development. A recent review of these roadmaps by Owen *et al*. [6] has confirmed that the internal combustion engine will continue to play a considerable role for the

[1] After-treatment devices are estimated to incur 8% penalty in fuel consumption because of the need for regeneration or through increased back pressure on the engine [3]. Other aspects such as safety developments have also affected engine fuel consumption through increased vehicle weight.

[2] Passenger car fleet averages of 130g/km to be phased in between 2012 and 2015. The long term proposal is 95g/km by 2020. Excess premiums will be enforced of up to €95 per exceeding g/km, per vehicle sold. There are currently similar regulations proposed for light duty commercial vehicles for 175g/km in 2012, reducing to 160g/km in 2015

foreseeable future. The review also states that "the impact of cumulative small improvements to vehicle system efficiencies had been much greater than expected". Longer term developments within the roadmap are dependent on significant technological breakthroughs, notably in the area of batteries. Shorter term improvements are expected through the "*pragmatic employment of known technologies*"[3].

For the internal combustion engine these pragmatic technologies relate to efficiency improvements that can be made with realistic cost impacts. The aim of improving the efficiency of the Internal Combustion (IC) engine is to maximise the work output per unit of fuel; studying the parasitic losses during engine warm-up falls into this area by reducing all auxiliary energy consumptions not associated with the direct production of useful work output.

2. *Parasitic losses*

For an internal combustion engine, parasitic losses relate to all the power consumptions that do not relate directly to useful work. These systems can be split into two groups: those that are vital for engine operation (lubrication, cooling) and those that are required for drive comfort/experience (power steering, air conditioning). The work in this thesis will focus on the first of these groups by assessing the lubrication and cooling systems.

2.1. Lubricant systems

Engine lubricant circuits are required to supply sufficient oil to all contact surfaces of the engine to avoid high wear rates that can ultimately lead to engine seizure. A typical production setup will consist of a fixed displacement pump driven directly from the engine crank. The oil will be filtered before being distributed to the working surfaces through a hydraulic network. In flowing through the engine structure, the oil often provides a cooling medium to the engine.

The lubricant circuit is designed to ensure safe engine operation in the most extreme conditions[4]. Consequently, oil systems are often over-specified and inefficient at other more

[3] Nick Owen of Ricardo UK, speaking at the IMechE low carbon vehicles conference, 20-21 May 2009, London, UK

[4] For the lubricant circuit, extreme conditions are when the lubricant has lowest viscosity and pump speed is low, i.e. idling under hot conditions.

common operating conditions. Oil systems influence engine fuel consumption in two ways: the power requirements from the oil pump and the friction at the contact surfaces.

2.2. Cooling systems

The engine cooling system must be designed to ensure that all engine metal and lubricant temperatures remain below critical thresholds under all operating regimes. Typical systems are driven by a mechanical centrifugal pump and a radiator allows the evacuation of excess heat. A wax element thermostat controls flow through the radiator and the cooling fluid[5] temperature. For modern engines, the coolant circuit is also required to cool exhaust gas recirculation (EGR) gases, oil and provide cabin heating.

Similar to the lubricant circuit, the cooling circuit is designed for worst case scenario[6]. Consequently at other more common conditions, the system is less efficient. The cooling system impacts on engine fuel consumption and emissions in two ways: the pumping work required and the control of thermal state, notably the impact on oil and combustion chamber temperatures.

2.3. Recent applications

Recent publications from major vehicle manufacturers at international conferences have shown significant interest in the topics introduced above and their uses on upcoming diesel and petrol engines. It should be noted that these developments have tended to be limited to higher end applications. Advanced cooling systems were presented using switchable coolant pumps, split cooling circuits [7-11] and active control of coolant flows [12, 13]. Friction reduction has also received considerable attention with the use of variable flow oil pumps [7, 10, 12, 13].

[5] The cooling fluid is usually a 50/50 mix of water and antifreeze to give good heat capacity for heat evacuation and continued operation under cold ambient conditions.

[6] For the cooling circuit, the most extreme conditions are when engine power is high but radiator cooling power is low, for example an up-hill trailer tow in hot ambient conditions.

3. *Aim and objectives*

The principle aim of this thesis is to demonstrate potential benefits to Diesel engine efficiency through pragmatic developments to the engine thermal and lubrication systems. Both of these topics could form the basis of a thesis in their own right, but this work will use a systems based approach to identify interactions and maximise benefits. A contribution to dynamic modelling will also be made, aiming at enhancing this methodology to capture engine behaviour over thermal transients.

The following objectives were laid out as a logical sequence, each contributing to the overall project aim:

1. Review the literature relating to engine cooling, lubrication and combustion focusing on interactions between the systems.

2. Establish experimental methods (hardware and procedures) to confidently demonstrate small changes in fuel consumption to prove cost effectiveness of candidate prototype hardware.

3. Oversee the integration of prototype hardware into a production engine.

4. Develop an understanding of the fundamental engine behaviour with changes to thermal system control during engine warm-up.

5. Develop an understanding of the fundamental engine behaviour with changes to lubricant system control.

6. Using the knowledge acquired in objectives (4) and (5), conduct an experiment to provide a detailed understanding of the effects and interactions of the prototype hardware. Use this to subsequently optimise the system calibration during warm up.

7. Demonstrate suitable system behaviour under fully warm conditions.

8. Improve state of the art engine dynamic modelling methods to capture engine behaviour during warm-up.

4. *Scope of thesis*

The main work of this thesis is organised into chapters 2 to 9. A 10th chapter brings together the conclusions and recommendations for the use of this work and additional research. An overview of the contents of the following chapters will be given in the next paragraphs.

Chapter 2 reviews the state of the art in the development of thermal management and lubrication systems. Novel concepts and the physical processes are presented and discussed. A brief overview of the diesel combustion system is then presented before the interactions between these three systems are critically assessed.

Chapter 3 reviews, assesses and proposes improvements to experimental procedures for the measurement of fuel consumption and emissions. Established and new model based techniques are presented that combine design of experiments and system identification theory. The measurement system for fuel consumption is analysed and preventative, corrective and procedural methods are proposed to improve the accuracy. Some aspects relating to testing repeatability are discussed before the capability of the experimental facility is measured.

Chapter 4 describes the experimental facilities and engine hardware used during this project. Candidate devices are then presented along with their integration into the production engine setup. Finally the instrumentation, data collection and data processing routines are presented.

Chapter 5 discusses the fundamental engine behaviour in response to variations of engine thermal state. In a first stage, engine friction is measured from cold-start to identify the relationship to operating conditions. Steady state tests are then used to understand the physical processes and interactions of the thermal system and the engine controller. Finally, transient tests assess how this behaviour manifests itself under drive cycle conditions using the candidate hardware solutions.

Chapter 6 analyses the behaviour of the engine in response to variations in the lubricant circuit. Local temperature measurement show the thermal effects of different external circuit configurations and varying oil flow.

Chapter 7 builds on the findings from chapters 5 and 6 to establish a detailed understanding of the interactions between the candidate hardware. A DoE approach is adopted to optimise the use of the candidate hardware for minimum fuel consumption following cold-start. The suitability of the optimised system for fully warm conditions is also discussed.

Chapter 8 attempts to improve the system calibrated in chapter 7 by the integration of an exhaust gas heat exchanger. The DoE approach is repeated for the new system, and results are compared to those from the previous chapter.

In **Chapter 9** the dynamic DoE approach is used to establish global dynamic models of the engine behaviour under fully warm conditions. This approach will subsequently be enhanced to capture engine behaviour during warm-up. The uses of these models will then be explained along with some example applications.

Chapter 10 summarises the process and findings from this project and bring a closure to the thesis. The aim and objectives will be considered against the achievements of the work to draw conclusions and perspectives from the project.

Chapter 2 - *Review of the interactions of thermal, lubricant and control systems*

In this chapter a review of the state of the art of engine thermal, lubricant and combustion systems is presented. Each of these systems will be reviewed independently to present the fundamental principles and developments in these fields. It is clear that both could be the focus of a thesis in their own right, and although these aspects are necessary for the purpose of this work, attention will be focused on the interactions between these systems.

In the same way, a brief review of the diesel combustion process, emissions formation and the common in-cylinder control mechanisms is also included. With this in mind, the interactions between the engine thermal management system and combustion process will also be discussed.

The analysis, discussion and conclusions from this chapter have been published in the form of a review article in the proceedings of the Institute of mechanical engineers part D: Journal of automobile engineering [14].

1. *Engine Thermal management system*

1.1. Current engine cooling systems

Thermal management systems (TMS) of modern Diesel engines have to perform a number of tasks around the vehicle in addition to maintaining integrity of the engine structure. The cooling system also has influences on the oil temperature with the use of oil coolers, the EGR gas temperature[7], safety attributes through the demist capability of the vehicle and the thermal comfort provided by the cabin heater [18, 19]. Developments presented in the literature can be split into the following categories:

- Reducing the power consumption to improve the efficiency of the system, i.e. reducing the pumping work required. In all but some of the most recent production engines, the coolant pump is directly driven by mechanical transmission from the crank shaft (belt, chain or gear). The pumping work required to circulate fluid around the system causes a load on the engine which ultimately increases engine fuel consumption.
- Optimising the use of the system by seeking benefits through interactions with other systems such as engine lubrication. Adjusting the operating temperature of the engine can be both beneficial to improved fuel consumption and improved thermal comfort by providing cabin heat and demist capabilities sooner. However, whilst the two are complementary, they can also be in conflict when designing heat distribution around the engine.

1.2. Improving cooling system efficiency

Conventional cooling systems are passive systems where a mechanical pump is directly driven from the crank shaft and the pumping power is directly linked to engine speed. A wax element thermostat reacts to top hose coolant temperature and distributes flow accordingly either through the radiator or straight back into the engine without significant cooling. The systems are typically designed to operate under extreme conditions (uphill trailer tow) and as such are over-specified in more common driving conditions. For example, in the case of motorway driving, the engine speed is high causing high coolant pump loads; however, due to the high

[7] Cooling EGR gases is essential for emissions control and to avoid reducing in-cylinder mass through thermal throttling [15-17].

speed of the vehicle, a large cooling power is available over the vehicle radiator [20]. This results in both excessive pumping work and overcooling of the engine. Consequently there is potential for improvements to the system both through improving heat transfer and using active systems to match cooling power to cooling requirements.

1.2.1. Active control of cooling circuit

Many examples of active control are presented in the literature; the most common being replacement of the mechanical pump and wax element thermostat with an electric pump and a control valve [20-22]. Other approaches have used coolant throttles in conjunction with the mechanical pump or the use of a thermostat that is also sensitive to bottom hose coolant temperature[8] [23]. These systems can reduce coolant flow rate by up to 90% [24] and may employ considerably downsized pumps (for example a 60W water pump may replace a 2kW mechanical pump [25]. The resultant change in fuel consumption is offset by increased power conversion losses associated with the electrical system, although a move to 42V systems would reduce these losses. There have been reports of fuel consumption gains of up to 5% [24], though these systems often change the engine operating temperature which also has an effect on fuel economy[9]. It is therefore difficult to isolate the benefit attributable to the improved system efficiency [19, 24, 27]. If the system can be used to reduce radiator fan on-time then this will contribute to lower fuel consumption [24]. The use of electronic components will increase cost and this is perhaps the main reason why they have rarely been put into production and only on more premium vehicles [7-11]. Recently, some high specification engines have been put into production using clutched water pumps as a compromise between the cost and benefits [7].

The active systems using electric pumps also allow continued coolant pumping after the engine is stopped. This helps avoid excessively high coolant temperatures during the thermal soak. Some authors [13] have designed a system using this principle which reduces the requirements on internal coolant volumes, thus reducing the overall thermal inertia, making engine warm-up times shorter.

[8] Cracking temperature was increased when bottom hose temperatures were low preventing overcooling and reduced when bottom hose temperatures were high (for example in trailer tow situation).

[9] The ability to actively control the thermal condition of the engine independently of engine speed also offers potential gains that can be exploited when incorporating the TMS into the control strategy. This will be discussed in the following section.

1.2.2. Enhanced heat transfer

A number of concepts have targeted improved heat transfer in the engines to reduce required coolant flow rates which in turn reduce the pumping power. Two of these concepts make use of the variation of heat transfer with local coolant velocity and temperature (see figure 2.1). Some of these approaches will be introduced below, however for a comprehensive review of these methods the reader is directed to the work by Pang and Brace [28].

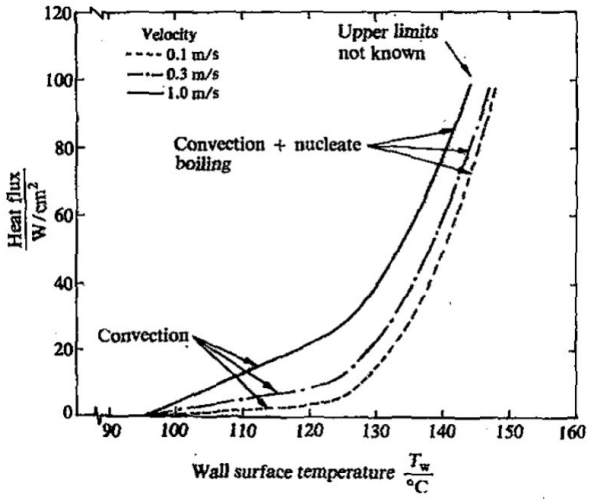


Figure 2.1: Heat flux variation dependent on coolant velocity illustrating effects of flow velocity and nucleate boiling [29]

Precision cooling: this technique was introduced by Clough [30] and focuses cooling inside the block and head more intensively to critical areas by adjusting passage cross-sections and introducing cooling jets into the flow. The work involves matching coolant velocities to the local heat fluxes to avoid excessive or insufficient cooling. Results from the original work reduced coolant flow by 65% however with the extensive use of computational models (CFD) for optimisation [31, 32] this philosophy is now embedded into mainstream engine design.

Nucleate Boiling: by allowing the fluid boundary layer to operate at its boiling temperature, the heat transfer coefficient is enhanced, making the system more efficient at removing heat. As the fluid near the surface begins to boil the heat flux is improved as illustrated by the change in gradients in figure 2.1. Small bubbles of vapour form at the convective surface. As the vapour moves into the colder bulk fluid, the bubbles quickly condensate back into a liquid state. Very few engines are designed for nucleate boiling and it is often used as a safety margin. It is important to note that if it does unintentionally occur, component temperatures could become insensitive to bulk coolant temperature.

Reverse cooling / Split cooling: The conventional coolant flow path in the engine is to enter the engine block before flowing into the cylinder head to allow easier vapour collection. Reverse or split cooling systems aim to supply colder coolant to the cylinder head where heat fluxes are higher and cooling more critical. Some recent production engines have been designed with split cooling circuits [8, 9, 11].

1.3. Modulation of engine thermal state

In passive systems, the wax element thermostat controls to a constant coolant temperature meaning at low load, the engine will operate significantly colder than at high load. The introduction of an active system offers the potential for more elaborate engine temperature control. The system could use a control mechanism to maintain a constant metal temperature over the engine operating map [33]. Increasing the engine temperature at part load would reduce oil viscosity and hydrodynamic friction, with the possibility of fuel consumption gains; this will be covered in more detail in section 3 of this chapter.

Coutouse and Gentile [27] developed a (TMS) control system around an electric coolant pump, electric control valve and shutters to control airflow over the radiator. Two coolant temperature set-points were defined based on engine load (115°C at low load and 100°C at high load). The advantage of the two set-points was to reduce friction at low load by allowing the oil to operate at a higher temperature whist avoiding excessive temperatures at high load. Similar control strategies have been implemented more recently both experimentally [34] and in simulation environments [26] but neither discuss any impacts on engine or oil durability.

The concept of variable coolant temperature was improved by Matteo *et al.* [33] who controlled coolant temperature to maintain a constant metal temperature over the engine operating map. This is also referred to as critical component temperature control (CCTC) [19] and the concept has recently been introduced in a production spark ignition engine [13]. An example of how the target coolant temperature is modulated over the engine operating range is shown in figure 2.2. Although in theory this approach offers fuel consumption benefits and these can be demonstrated experimentally under steady state operating conditions, the authors do not discuss the dynamic behaviour of the engine due to the large thermal inertias and response time when changing temperature set-point.

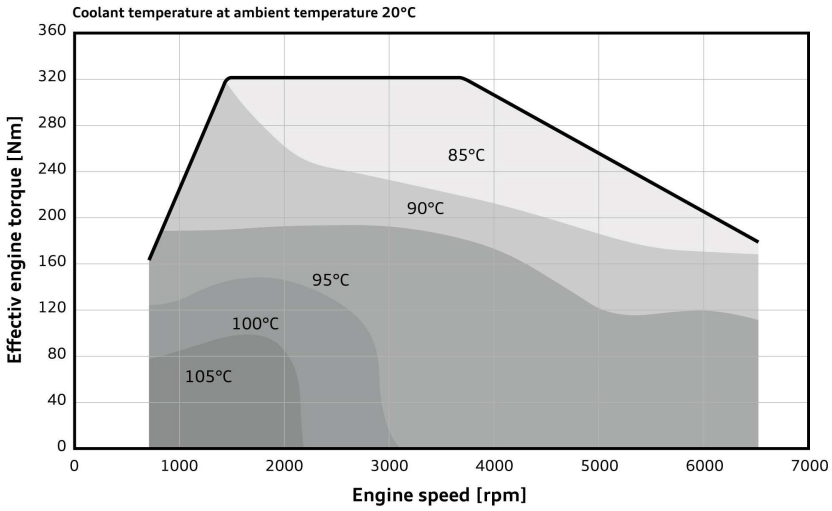


Figure 2.2: Example of coolant temperature modulation over the engine operating map to maintain more constant metal temperatures and reduce frictional losses at lower loads [13]

1.4. Improving engine warm-up

Modulating engine temperatures when running under fully warm conditions has clearly shown potential benefits in fuel consumption, however in many real world applications the engine will operate under transient thermal conditions following a cold-start[10]. During this time the available heat energy is scarce and the fuel consumption penalty is high so improving the warm-up behaviour presents clear benefits. Also, improvements in cold-start will improve both passenger comfort by faster cabin-heater performance and safety by earlier demist capabilities [19]. The effect of engine warm-up on vehicle fuel consumption has been quantified by Kunze *et al.* [36] who compared the fuel consumption over a cold-start and hot-start New European Drive Cycle (NEDC). The authors reported 10% higher fuel consumption over the cold test. Results from the University of Bath on an engine dynamometer show a 4% difference (cold-start at 25^{o}C). The additional penalty from the literature is a result of the additional powertrain elements and their cold-start contributions (gearbox, differentials and tyres).

A series of concepts have been tested aiming to improve warm-up time with contrasting results. Brace *et al.* [23] installed a throttle at the outlet of a conventional mechanical pump, but whilst steady state temperatures were 10 to 15^{o}C hotter with the throttle partially closed, the system had no effect on warm-up rate over an NEDC. Similar results were observed in

[10] Studies of real world usage of passenger cars have shown that about a quarter of journeys are less than 1km and over half less than 3km [35].

other systems [33, 37] (see figure 2.3), which suggest that reductions in coolant flow rate have minimal effects on warm-up rate, but significant effects on steady state operating temperature.

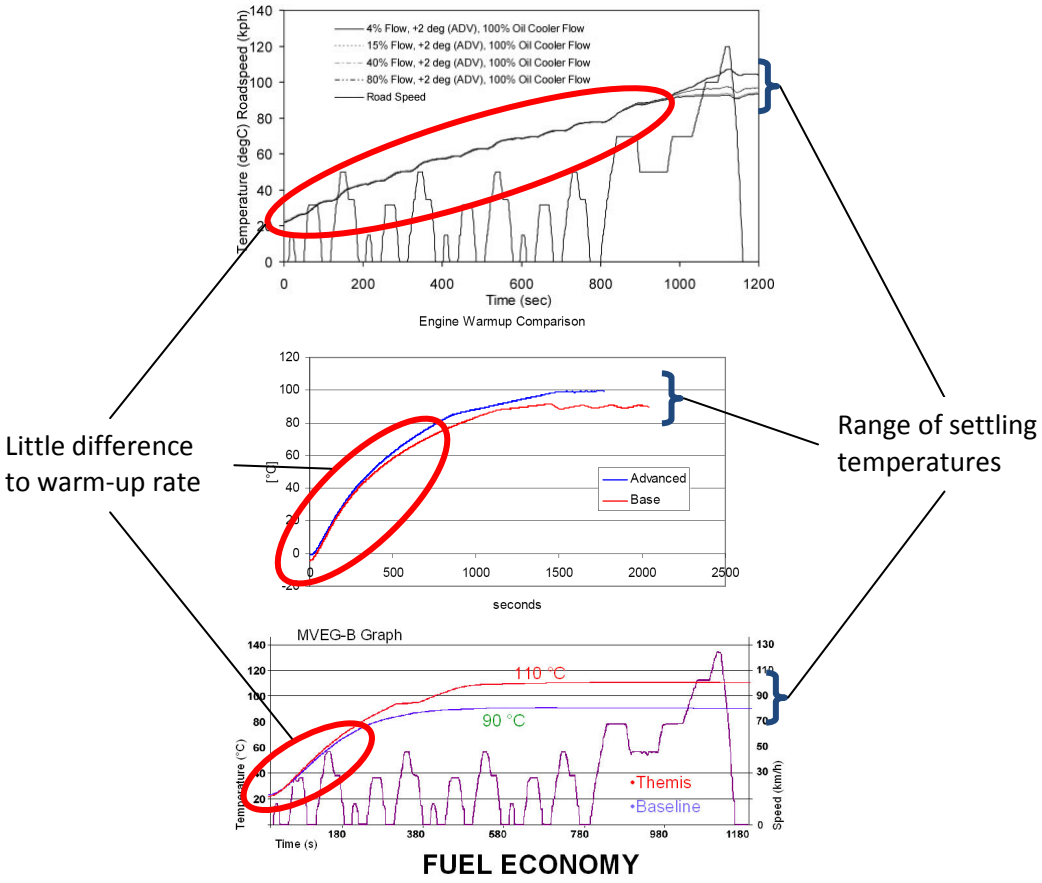


Figure 2.3: Warm-up rates using throttled coolant flow [23] or electric coolant pumps [22, 37]. Each have little effect on warm-up rate but do impact steady state temperatures

An interesting comparison to these studies is the work published by Choukroun and Chanfreau [34]. Using an electric coolant pump and control valve they limited coolant flow and, as with other results [23] achieved higher steady state operating points and reduced fuel consumption by 2% for a 20°C temperature increase. By completely stopping coolant flow in the engine over the first 300 seconds of a cold-start NEDC, they reduced the time for coolant to reach 100°C following a 25°C start from 600 to 300 seconds. Close examination of the coolant system shows that the control valves isolate a large section of the cooling circuit during warm-up, effectively reducing the thermal inertia of the system. This results in a 2-3% benefit in fuel consumption. Further experiments presented in their work consider a reduction in cooling flow in addition to isolating a portion of the circuit. This reduction in coolant flow offers little

further benefit and is effectively the same process shown in figure 2.3. Similar results were presented by Krause and Spies [38].

Reducing the thermal inertia has proved successful in reducing warm-up time [19]. Precision cooling systems have contributed to this by designing smaller passages within the engine. In the work by Clough [30] coolant jacket volume was reduced by 64%, reducing warm-up time by 18%. However unless inertia is reduced temporarily using flow control valves, experimental work is limited because reductions are only achieved through significant engine design changes. As a result, model based approaches tend to be used based on *thermal lumped capacity nodal models*. These models are constructed by linking simulated thermal masses through different modes of heat transfer and different examples are available in the literature [39-42]. The main difficulty is validating each of these masses and heat flows to provide realistic predictions. Using such a model, Torregrosa *et al.* [43] simulated the effect of reducing coolant volume and reducing flow rate and found that reducing the mass of coolant was the most significant effect to shorten warm-up time, effectively reaching the same conclusions as Choukroun and Chanfreau [34]. The concept was then produced experimentally and the combination of reduced coolant volume and flow reduced fuel consumption by 1.64%.

Faster engine warm-up can also be achieved by additional heat injection into the system however concepts that involve additional energy usage[11] will be detrimental to fuel consumption [44]. Energy recovery has a large potential for improved warm-up, however designing a system that is subsequently capable of rejecting the excess heat under fully warm conditions is problematic [19]. Possible ways to recover otherwise waste energy include coolant to exhaust gas heat exchangers [45, 46] or the storage of energy from previous warm operation by use of insulation [47] or heat storage devices[48, 49].

2. *Engine Lubrication system*

The benefits from improved warm-up and hotter engine operation are always simply explained by lower engine friction as a result of changes in oil viscosity [18, 23, 28, 50]. It is therefore necessary to review engine friction behaviour. Initially a description of engine friction and lubricant chemistry will be presented. Secondly improvements to the efficiency of the system

[11]To meet cabin comfort requirements, forcing the engine to work harder for example through increased battery charging causes an increase in fuel consumption but achieves faster warm-up rate. Fired cabin heaters are also used for this effect.

will be reviewed followed by the thermal effects on engine friction. Finally further considerations for engine warm-up will be discussed.

2.1. Engine friction basics and measurement

Friction in internal combustion engines is strongly linked to the conditions experienced by the lubricating oil. As a result, depending on speed, load and oil properties three regimes can occur (see figure 2.4) [51]:

- **Hydrodynamic lubrication**: an oil film builds up between the two surfaces and completely separates them. In this case the friction is only a result of the shearing of the fluid and highly dependent on the relative speed of the two surfaces and oil viscosity.
- **Mixed lubrication regime**: As the oil film breaks down due to reduced speed or reduced oil viscosity, some contact occurs between the two surfaces and friction can rise quite rapidly. It is important to note that reducing engine friction cannot be achieved by reducing oil viscosity indefinitely.
- **Boundary Lubrication**: the surfaces are in contact and rubbing against each other and friction is high and depends on the strongly on the normal load and surface properties of the two materials.

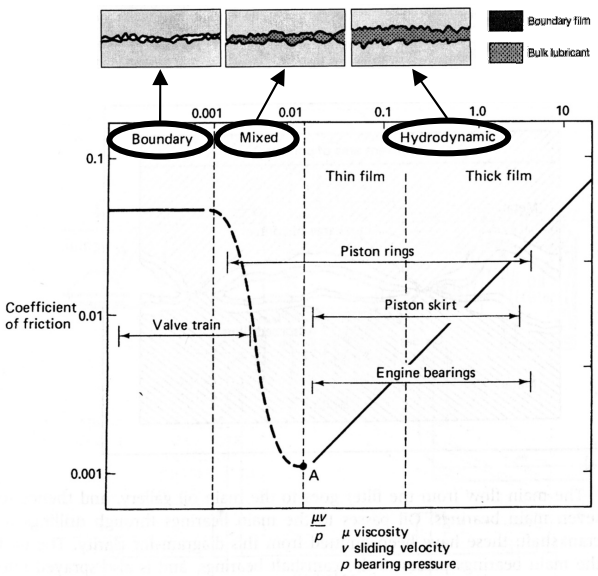


Figure 2.4: Stribeck curve showing different lubrication regimes and operating points of key engine components. Adapted from [51] and [52]

Experimentally, only one technique exists to measure total engine friction in a firing engine: the indicator diagram method [52, 53]. Indicated mean effective pressure (IMEP) is compared to brake mean effective pressure (BMEP)[12]; the difference is the friction mean effective pressure (FMEP), which will include the load from auxiliary devices. The issues with this method involve the correct alignment of the pressure signal that depends heavily on being able to accurately detect top dead centre (TDC) [54]. Also, as both BMEP and IMEP are large compared to FMEP, the latter is ill-conditioned and difficult to distinguish accurately from the errors of the two larger measurements. In addition, this method does not give any insight into the contribution of individual components to total engine friction [55, 56]. Other measurement methods involve motoring the engine, or specific parts of the engine (teardown tests) but fail to represent real operating conditions due to lower temperatures and reduced in-cylinder pressure [57].

Every rubbing surface within the engine will contribute to mechanical friction to an extent [58, 59], but the components studied by the majority of authors are the piston assembly, the main bearings and the valve train. Figure 2.5 shows the scatter of results, but piston assembly contributions tend to represent 40%-70%, main bearings 12%-40% and valve train 19%-35% [53, 60]. The engine to engine variations will be a result of design (more hydrodynamic or boundary friction), lubricant chemistry (including grade and additives) and test operating conditions [61].

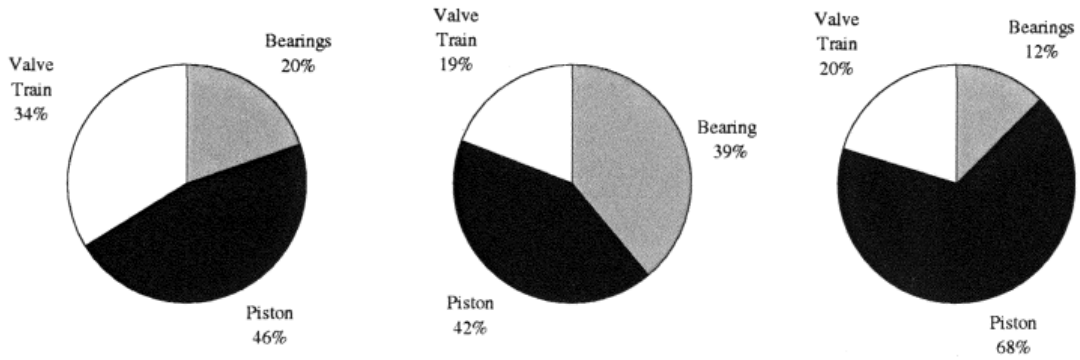


Figure 2.5: Friction breakdown for various engines for piston, bearings and valve train. These show the clear engine to engine variations in friction levels that are a result of engine size and design [60].

[12] IMEP is typically calculated from in-cylinder pressure data while BMEP is calculated from engine speed and torque measurements from the dynamometer.

Friction in the piston assembly is the most complex as large variations in speed and load occur over a single cycle. The friction is the result of both the piston rings and the piston skirt rubbing against the cylinder liner. The rings are loaded against the liner through their pre-tension and under the effect of the in-cylinder pressure. The skirt will tend to lean against the liner as a result of the reaction of the conrod. It is generally assumed that the piston skirt acts in the hydrodynamic region throughout the engine cycle, whilst the piston rings have a more complex operation [62-65]. When the piston is mid-stroke, hydrodynamic lubrication occurs. As the piston approaches top or bottom dead centre, the reduction in speed causes the oil film to break down[13]. At TDC combustion, it is enhanced by the high in-cylinder pressure which increases the loading further [53, 54]. Overall, piston friction occurs mainly in the hydrodynamic regime and is therefore highly sensitive to the lubricant viscosity.

The main and big-end bearings operate in the hydrodynamic regime, though there are claims of mixed lubrication during peak cylinder pressures [66]. As for piston assembly, bearing friction is sensitive to oil viscosity.

Valve train studies are limited but suggest boundary friction dominates between the cam and follower. This has been observed following reports of increased friction with engine load [54, 67, 68] and reduced friction with increasing engine speed [54][14]. As the valve train does not operate in the hydrodynamic regime, it is less sensitive to oil viscosity and can even increase for reduced viscosity [61]. However, boundary friction can be influenced by friction modifier additives to the lubricant.

2.2. Lubricant Chemistry

Lubricants are the result of a base stock which can be of either mineral or organic origin and a collection of additives which represent only a small fraction of the final product [69]. Automotive engine oil viscosities are highly dependent on temperature and reduce exponentially with increasing temperature [51]. This relationship is quantified by the *viscosity index* (VI) which is an arbitrary scale assessing the change in viscosity between 38^{o}C and 100^{o}C. Initially the scale of VI was 0-100, though to perform satisfactorily in modern engines, oils now

[13] As speed reduces at constant load and viscosity, there is a shift to the left on the Stribeck curve (figure 2.4)

[14] An increase in engine speed causes a shift to the right on the Stribeck curve meaning a resultant decrease in friction would suggest moving from boundary to mixed lubrication regime.

have VI levels above 150 [51, 69]. To achieve this, additives known as *viscosity modifiers* or *Viscosity Index Improvers* are used. These are polymeric molecules that have a temperature-dependent structure: at low temperatures they coil into a ball and have little effect on the fluid viscosity, but at high temperatures they uncoil and considerably increase the oil viscosity [69]. Figure 2.6 shows the viscosity temperature relationships for two single-grade oils and one multi-grade oil: both single grade oils have VI of 95 and their viscosities reduce similarly with temperature. Society of Automotive Engineers (SAE) grade 5W-30 reduces a lot less with temperature and has a VI of 162. The multi-grade oils would be obtained from the base stock of the SAE 5W oil with the addition of the appropriate concentration of friction modifiers to obtain the viscosity of SAE 30 grade oil at higher temperatures.

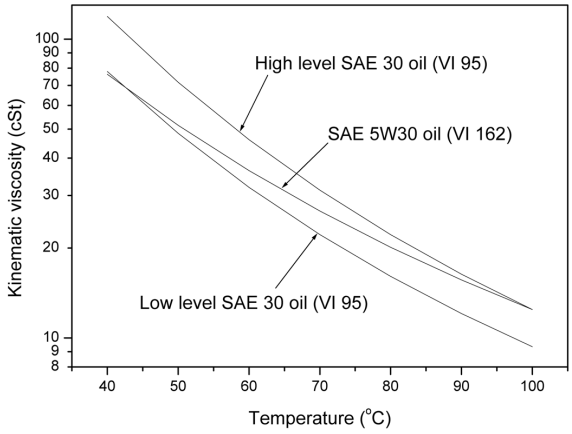


Figure 2.6: The effect of Viscosity modifiers on base stocks, showing how viscosity modifiers change the temperature response of a base oil to produce multi-grade oils, adapted from [70] (note the logarithmic scale of y axis)

Other important additives are friction modifiers (FM), designed to improve the boundary lubrication characteristics, reducing the friction coefficient at the very left hand section of the Stribeck curve (figure 2.4). These molecules have a polar constituent that attaches to the lubricated surface, whilst organic chains in the molecule absorb a layer of oil. They are very effective at reducing friction in boundary lubrication as long as the temperature does not rise such as to cause decomposition of the molecule or desorption of oil on the surface [69].

The effects of FM have been tested both on engine [71] and isolated test rigs, such as a Cameron-Plint machine[15] [72] and give inconsistent results. Two types of FM were tested, Molybdenum (MoDTC) and an organic FM and at two different concentration levels. On the Cameron-Plint machine, the MoDTC in high concentration yielded the largest reduction in

[15] The Cameron-Plint machine is used to simulate piston liner friction.

friction coefficient (66%) whilst the reduction for the organic FM was only 16%. In contrast, during engine dynamometer testing the organic FM offered the best FMEP reduction (4.5%) whereas the MoDTC showed no significant improvements. The differences were suggested to be caused by different operating conditions (loading, temperatures, soot contamination…). However this highlights the difficulty in performing accurate testing in this area.

Other additives are necessary in engine lubricating oils such as antioxidants, detergents and pour point depressants and for a comprehensive understanding the reader is directed to other published work [51, 69, 73].

2.3. Improving lubricant system efficiency

The common production oil circuit usually incorporates a fixed displacement pump which is specified to supply sufficient oil flow and pressure in the most extreme conditions[16]. At higher engine speeds and oil viscosities, a pressure relief valve is installed to limit maximum oil pressure which could be dangerous for engine operation. The conventional oil system design is a reliable and cost-effective solution, however there is clear energy waste when using the pressure relief valve. The upper part of figure 2.7 shows some typical lubricant pressure requirements with engine speed and the black line shows typical oil pressure supplied to the engine. The authors of that work did not explain the location of the *main engine hole,* but it is assumed this is representative of engine main supply or oil gallery. The requirements of key components are shown along with the area where the pressure relief valve is required for engine speeds above 4000rpm. If the oil pump supplies a higher pressure than is required it will correspond to an energy waste.

A number of authors have modulated the oil flow according to engine requirements by using a variable flow oil pump (VFOP) [74-79]. The simplest approach consists of reducing the displacement volume at higher engine speeds to avoid oil flow through the relief valve but Toyoda *et al*. [76] also reduce pump work at lower speeds (as in figure 2.7). The lower part of figure 2.7 shows the corresponding reductions in oil pump driving torque.

Rundo and Squarcini [74] compared a vane type VFOP with an external gear and gerotor fixed displacement pumps. The project was quite conservative as it concentrated on exploiting the

[16] It was seen earlier this corresponds to hot, low viscosity oil and low engine speed.

VFOP only during warm-up to reduce oil pressure with cold lubricant. This showed fuel consumption benefits of 0.5% during warm-up over the NEDC, but for longer journeys the VFOP was detrimental due to its reduced hydraulic efficiency. The approach by Brace *et al.* [74] was more aggressive in that oil pressures were significantly reduced over the engine speed range compared to the production pump. Two designs were compared, the first an internal rotor and the second a vane design and yielded 2.6% and 3.4% fuel consumption (FC) benefits respectively. These benefits have clearly been recognised by manufacturers and recent high end production engines are incorporating these pumps [7, 10, 12, 13].

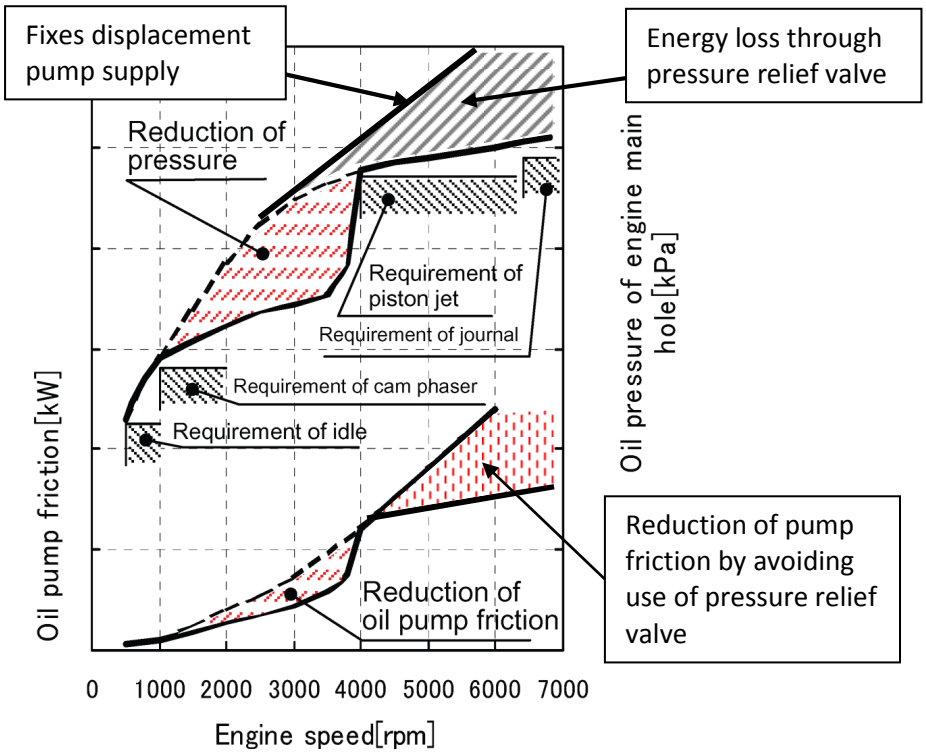
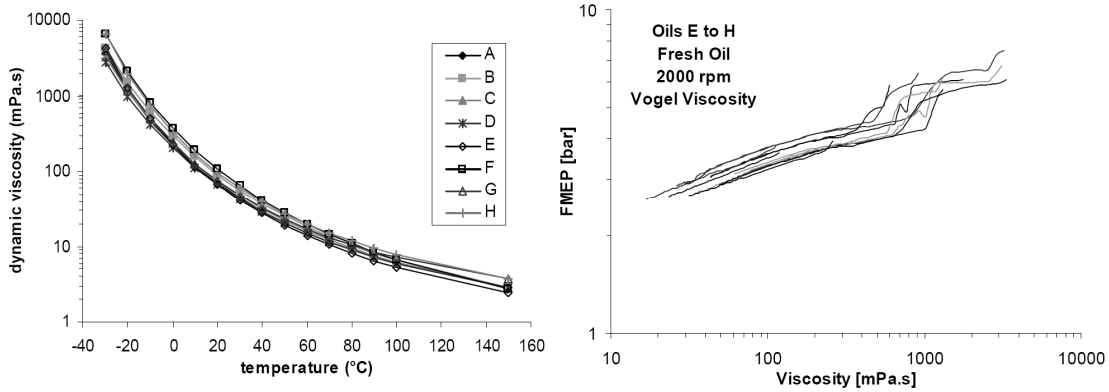


Figure 2.7: Optimised oil pump usage over low speed region: Shows oil pressure requirements along with fixed displacement pump supply to highlight possible reductions in pumping power (adapted from [76])

Most previous studies on the use of VFOP on engines have concentrated on the fuel consumption effect as a direct result of reduced pumping work. However, with the inclusion of piston cooling jets and oil coolant heat exchangers, the interactions with the thermal behaviour of the engine are also important. Agarwal and Varghese [80, 81] have published considerable work analysing the cooling effects of piston jets. Using finite element simulation piston cooling jest were predicted to reduce piston temperatures by around 40^{o}C under fully warm conditions. Their work has also highlighted the impact of jet velocity on piston temperatures. It is important to note that jet velocities will be strongly linked to changes in oil pressure.

2.4. Thermal state and engine friction

Shayler *et al.* [82] show the overall effect of engine temperature on engine friction (figure 2.8). Increasing oil temperature causes an exponential drop in oil viscosity which causes a reduction in overall engine friction. In reality this reduction would be a cumulative effect of changes in friction throughout the engine. In all cases, when oil viscosity is reduced, a shift from right to left on the Stribeck curve occurs (figure 2.4). Where friction occurs in the hydrodynamic regime, a reduction in oil viscosity will reduce engine friction. In contrast, areas where friction occurs in the mixed regime will see rises in friction for reduced viscosity. Overall, the hydrodynamic lubrication regime appears dominant. This is consistent with various contributions to total friction described previously.



(a) *(b)*

Figure 2.8: Dynamic viscosity (a) and measured engine FMEP (b): as oil temperature increases, oil viscosity decreases which decreases FMEP [82]. Traces A to H represent different oils

Wakuri *et al*. [54] measured engine friction using the run-out method[17] on an indirect Diesel engine. The authors vary the cooling water temperature under steady state conditions to vary the oil sump temperature. At a speed of 2000rpm they see a 44% drop in engine friction when oil temperature is increased from 50°C to 80°C (see figure 2.9).

Daniels and Braun [55] studied the individual component contributions with coolant temperatures of 25°C and 85°C using motored teardown tests. The variation in coolant temperature was reflected in the sump oil temperature (25°C to 77°C). The results showed significant reductions in piston assembly and main bearing friction, 66% and 85% respectively,

[17] The run out method consists of running the engine to steady state operating conditions and then suspending firing. The deceleration of the engine is then studied to determine FMEP. Whilst this method fails to include the effect of in-cylinder pressure which plays a significant part in both piston assembly and bearing lubrication, it does allow for representative oil temperatures.

but an increase in valve train friction (33%). This can be explained by referring to the Stribeck curve (figure 2.4) and confirms that the total change in friction following oil viscosity change is the result of opposing friction increases and decreases at different locations in the engine. The valve train operates in the mixed and boundary lubrication regime and as the engine oil warms up, its viscosity reduces pushing the operation further into the boundary regime. It should also be noted that changes in thermal conditions may affect the performance of boundary friction modifiers. In contrast, the piston and bearings operate in the hydrodynamic regime and their friction reduces with oil viscosity. Due to the relative friction contributions of each component group, the overall FMEP was seen to reduce by 26% (see figure 2.10). This is significantly lower than the results published by Wakuri *et al.* [54] which may be explained by different engine designs and operating conditions.

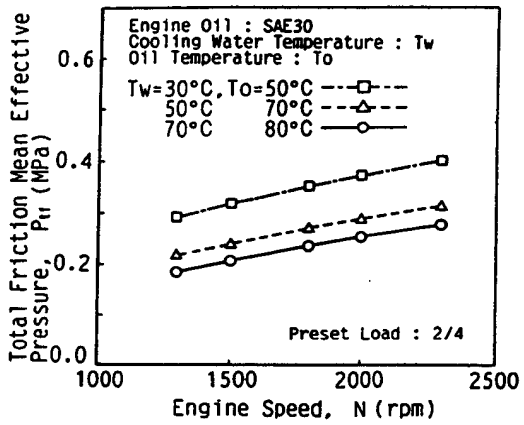


Figure 2.9: Effect of engine speed and coolant temperature on total FMEP [54]

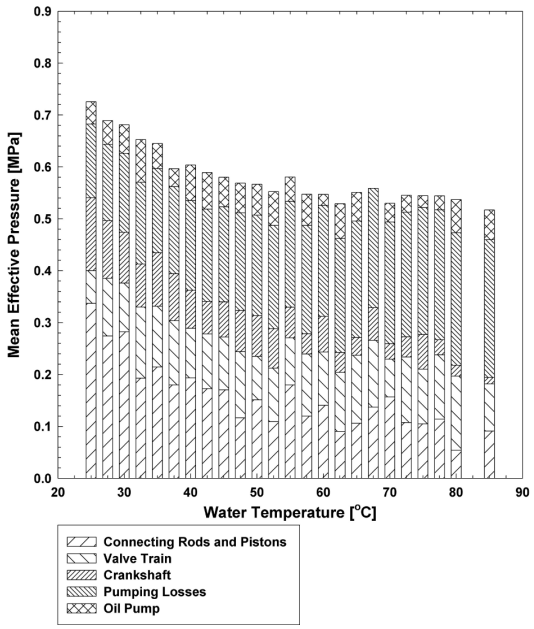


Figure 2.10: The effect of coolant temperature on friction from individual engine components[55]

2.5. Engine friction models

A large number of engine friction models have been published with various levels of sophistication [60-64]. The simpler models are often related to engine speed and use assumptions based on the lubrication regimes of individual components. More complex models attempt to determine oil film thickness by solving the Reynolds equation and are based on assumptions of inlet and outlet conditions in precise locations in the engine. Taylor [61], using a model based on solving the Reynolds equation [83], predicted a 66% reduction in friction when oil temperature is increased from 30^{o}C to operating temperature (between 100 and 150^{o}C depending on components).

The majority of engine friction models are very crude in terms of engine warm-up simulation and generally have oil temperature as a model parameter taken from measurements. Whilst engine friction models are often used to determine the friction heat generation as an input to the lumped capacity models (see section 2), the results from the lumped capacity models are rarely fed back into the friction models to assess full warm-up capabilities. One example is the work by Jarrier *et al.* [39]; the authors used a nodal model in conjunction with an empirical friction model to assess the effect of reducing oil volume from 5L to 1L. This yielded a predicted 7^{o}C higher temperature during warm-up over NEDC, resulting in a 5% drop in FMEP. Whilst these results show a similar trend to the experimental work by Law *et al*. [84], the use of an empirical friction model limits analysis to very specific applications.

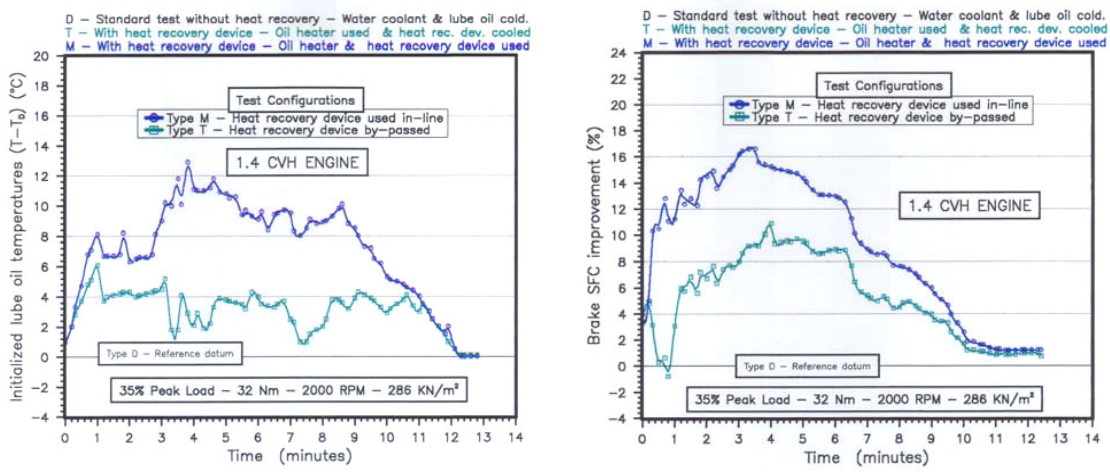
3. *Using the TMS for lubricant warm-up*

Experimental studies on engine where the warm-up rate of oil is specifically targeted are limited. The work by Brace *et al.* [23] has already been mentioned where the coolant flow rate was limited to minimise heat loss from the engine but had little effect on warm-up rate. When assessing oil temperatures, as with coolant temperatures, the authors found little difference in oil temperature over the range of coolant flow rates.

A second approach is the use of heat addition into the oil. Kunze *et al.* [36] used their model to assess the effect of increased heat addition into various fluids during NEDC cycle. They predicted a 1.5% fuel consumption benefit by adding 2MJ of heat to the oil, but they offer no insight into how this may be implemented practically. In addition, their predictions were not reproduced experimentally, where an increase in fuel consumption was obtained from

preheating the engine fluids. It was not clear whether this was due to a lack of precision in the model or the constraints of experimental methods, but does highlight the difficulties in transferring results from the nodal models to the real engine. Since engine friction is difficult to measure, mathematical models suffer in terms of confidence and accuracy. The engine thermal models that subsequently use these friction models inherit these inaccuracies.

The work by Andrews *et al.* [46] is a practical implementation of the study by Kunze *et al.* [36]. Heat was added to the coolant during the warm-up period by installing a coolant heat exchanger in the exhaust manifold. The heat was then transferred to oil via another heat exchanger. The tests were run from cold start on a steady state rig. The additional heat increased the warm-up rate of the oil by an average of 8 to 12^{o}C, yielding a 12-15% reduction in instantaneous fuel consumption as shown in figure 2.11. The figure also shows the effect of a coolant/oil heat exchanger alone. It can be seen that this provides a significant improvement in fuel consumption (8-10%) during warm-up. Despite the benefits in this study, the authors did not consider knock on effects on the after-treatment devices which require heat from exhaust gases to ensure tailpipe emissions remain within legislated limits.



(a) *(b)*

Figure 2.11: Measured oil temperature differences (a) and specific fuel consumption (b) compared to a baseline configuration using two levels of heat recovery during warm-up[46]

As was seen in the case of coolant circuit, warm-up rate of the oil may be improved by the reduction of oil volume. As can be seen in figure 2.12, bulk oil temperature lags behind coolant temperature and remains almost constant following cold start for the first 150 seconds. This is because the bulk oil in the sump acts as a heat sink to the returning oil that has been heated as it was fed through the engine. Obviously, reducing the thermal inertia of the sump oil will improve warm-up rates [46].

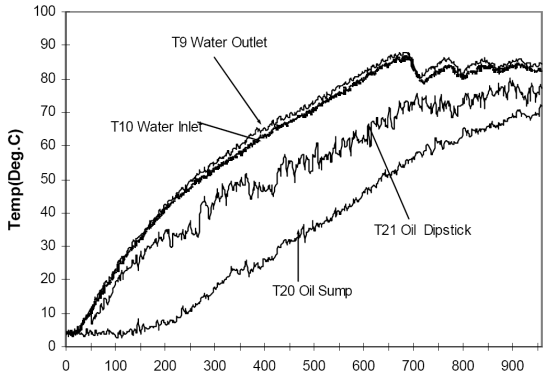


Figure 2.12: Coolant and oil warm-up rates over NEDC [46]

The conflict of reducing oil volume is the need to maintain reasonable service intervals. As a result, the total volume of oil should not be reduced significantly. Law *et al.* [84] used novel sump designs based on wire mesh stratification or the use of additional thermostat to encourage the circulation of a smaller volume of oil during warm-up. When the oil was hot, the total volume of oil was used to ensure oil durability. The improvements in fuel consumption as a result of the new system are dependent on engine start temperature, as for the case of overall engine warm-up (section 2). From a -10°C start, oil temperature was found to be up to 15°C hotter over the first two minutes and this resulted in a 10% reduction in engine friction (determined using the indicator method).

Previous studies have highlighted the thermal link between engine friction and thermal management systems. These interactions may be used to improve fuel consumption through reductions in engine friction with considerable gains following cold start. The absolute value of improvement is a complex result of opposing changes in friction in different parts of the engine. It is difficult to directly compare results from different studies because of the differences in engine design but also in test methods. Gains are often significantly higher if the engine starts from a colder temperature.

4. *Direct Injection Diesel combustion*

4.1. Diesel combustion process

In direct-injection Diesel engines, combustion occurs when high pressure fuel is sprayed into the cylinder at the end of the compression stroke. As the fuel burns, the rate of energy release will take a form shown in figure 2.13 [73]. The combustion event can be split into four phases:

- **Ignition delay period**: This is the time between the start of injection and the start of combustion. During this time liquid fuel is being injected into the cylinder and is being heated by the surrounding gases to its self-ignition temperature.
- **Premixed combustion phase**: This is an initial, violent combustion when the fuel that has evaporated during the ignition delay combusts rapidly. This causes a spike of heat release and is also the main reason behind the loud noise from diesel engines. The premix spike is also responsible for a large and rapid temperature rise that has a large influence on NO_x emissions.
- **Mixing-controlled combustion phase**: At this stage, the combustion is controlled by the rate at which the fuel and air mix in the cylinder.
- **Late combustion phase**: as the piston expands and the temperature drops, the final drops of fuel will finish burning and a chemical quenching will occur in the cylinder stopping the oxidation reactions of CO and HC.

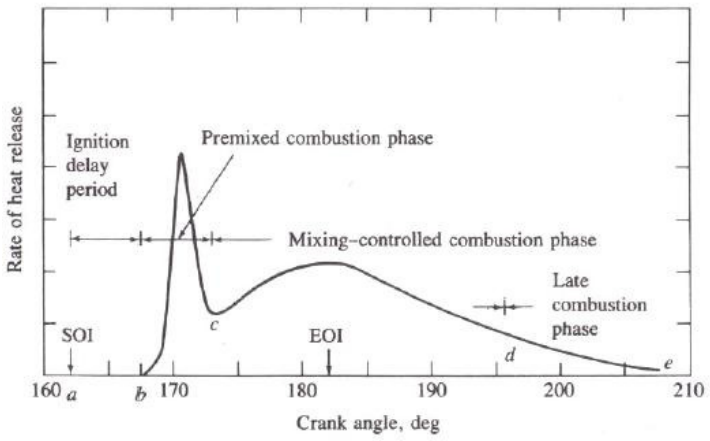


Figure 2.13: Typical heat release curve for Diesel combustion (SOI; Start of injection; EOI: End of injection) [73]

4.2. Emissions formation

The main products of combustion of diesel fuel are CO_2, water, NO_x, CO, HC and soot. CO_2 and water are natural products of the reaction and reducing CO_2 emissions is achieved by reducing the overall vehicle fuel consumption. The other emissions are by-products resulting from the specific conditions during and after the combustion event.

4.2.1. NO_x formation

NO_x emissions are primarily composed of NO and form according to four mechanisms described by Musculus [85]:

- Thermal NO
- prompt NO
- fuel NO
- Dissociation of N_2O

In classic Diesel combustion, thermal NO represents is the largest contributor [86]. The reaction is described by the Zeldovitch mechanism [73] and the speed of reaction is strongly dependent on temperature. For the reaction to have a significant impact on the time scales of Diesel combustion, the ambient temperature needs to be above 2000K [73, 86, 87], however NO subsequently increases exponentially with temperature. Control parameters that are used for NO_x control are measures that tend to reduce the combustion temperatures.

4.2.2. Products of incomplete combustion (CO, HC and soot)

CO, HC and soot are formed throughout the combustion phase within the Diesel jet [86]. They are the products of the rich premixed flame and serve as reactants for the diffusion reaction. Their formation is mainly linked to the quenching of the combustion gases as the combustion chamber expands. After the reactions have frozen, some of the molecules will then group together to form agglomerates that form particulate matter [88] (see figure 2.14). Measures that are applied to control NO_x emissions by reducing combustion temperatures tend to have an opposite effect on these emissions.

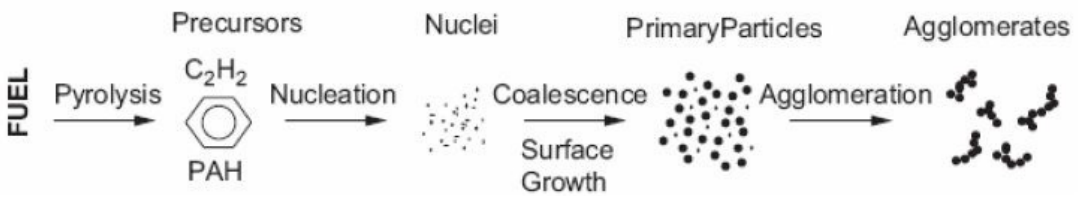


Figure 2.14: Soot formation process from fuel to agglomeration into PM [88]

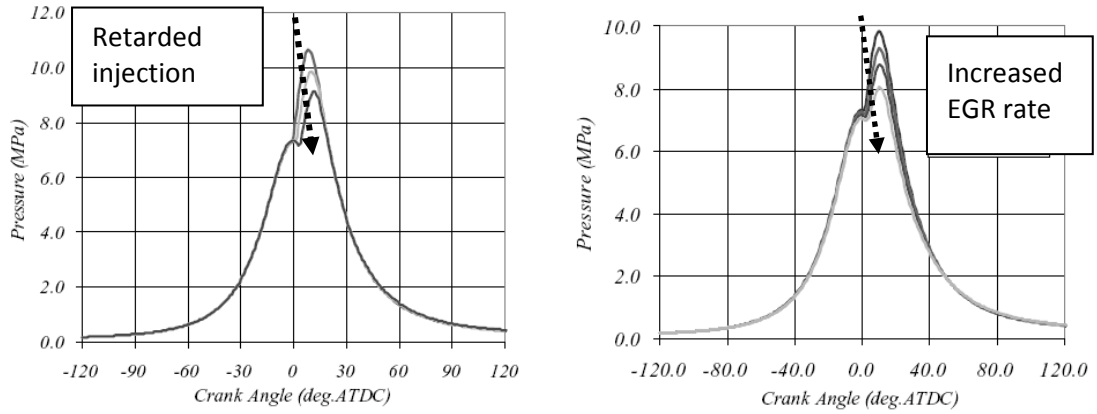
4.3. Common combustion control parameters

A number of common parameters have been developed to control the combustion process and allow the calibration engineer to trade off different emissions. Four of these will briefly be discussed here: inlet charge cooling, injection timing, injection rate shaping and EGR.

Cooling inlet charge gases is achieved practically on most turbocharged engines using an air cooled intercooler. By reducing the air temperature entering the engine, the temperatures are reduced throughout the whole combustion cycle thus reducing NO_x and increasing HC [89]. When inlet manifold gas temperatures were reduced from 70^{o}C to 45^{o}C NO_x reduced 20-30% and HC increased 30-40%. Other authors have identified an optimum inlet temperature for NO_x reduction as if temperatures are reduced too far, then ignition delay will increase causing a higher premix spike [85].

By adjusting injection timing, the entire combustion process can be shifted later or earlier in the cycle. If combustion occurs later then the piston will be further into the power stroke and as a result the peak pressure and temperature will be lower (see figure 2.15 (a)). This has a positive effect on reducing NO_x emissions and causes an increase in soot [90]. Fuel consumption is also affected as lower in-cylinder pressures will reduce the indicated work. The opposite effect is seen when injection is advanced.

Injection rate shaping aims to reduce the amount of fuel injected into the cylinder before ignition. It has only been made possible through recent developments of common rail injection systems with faster response times. Initially this was achieved using a small pilot injection before the main injection which raised the temperature in the cylinder to shorten ignition delay. The latest generation use boot-shaped injections where injection rate is slower over the first part until the jet is fully developed [91].



(a) *(b)*

Figure 2.15: Changes in cylinder pressure at different (a) injection timings and (b) EGR rates [90]

Exhaust gas recirculation (EGR) acts by diluting the inlet charge with exhaust gases and can reduce NO_x emissions at the expense of increasing PM and fuel consumption. Ladommatos *et al.* [92] showed that EGR acts according to three separate principles. The dilution and thermal principles are of particular interest in this study and account for the majority of the effect of EGR. The dilution effect limits the availability of oxygen thus reducing the formation of NO_x. The thermal effect increases the thermal capacity of the charge by introducing tri-atomic gases (water and CO_2) in the place of diatomic gases (nitrogen (N_2) and oxygen (O_2)). This increased thermal capacity reduces average in-cylinder temperature and pressures (see figure 2.15 (b)). The dilution effect also increases the amount of air entrained in the combustion jet due to the lower concentration in O_2: this increases the thermal capacity of the jet itself and reduces peak temperatures.

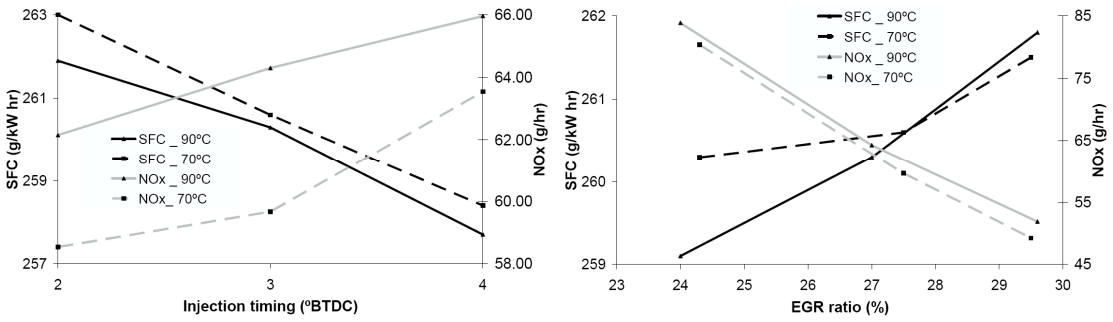
It is important to note at this point that the engine control unit (ECU) controls these parameters based on the engine operating conditions and that engine temperature is a critical input to this control system. For example, at lower temperatures, injection tends to be advanced so that fuel is injected when the air is at a higher temperature. This ensures that the fuel can evaporate and combust and avoids misfire. At higher temperatures, injection is retarded with the aim of reducing peak in-cylinder temperatures for NO_x control. As a result, any effects seen when influencing the engine TMS must be balanced against the reaction of the control system. For example, changes in fuel consumption and emissions may be observed as a result of changing the thermal state of the engine, but may be partly a result of a different operating strategy (injection timing, injected volume demand, EGR rate…) rather than physical influences of the changed temperature.

5. *Interactions of the TMS and combustion process*

5.1. Evidence of physical behaviour

Figure 2.16 shows results presented by Pang *et al.* [93] showing NO_x/FC trade-offs for two different operating temperatures along injection timing and EGR swings. From figure 2.16 (a) it can be seen that as timing is retarded, the effect of coolant temperature on NO_x becomes less significant. Both injection timing and thermal condition have similar effects on NO_x, but injection timing seems to affect fuel consumption more significantly. At advanced timings the effect of a 20°C drop in coolant temperature is to reduce NO_x by 3.5g/h (5.6%) whereas at more retarded timings the reduction is only 2.5g/h (3.7%). Over this same range, fuel consumption increases by about 1g/kWh (0.4%). It is unclear from the data whether the curvature in fuel consumption change and maximum in NO_x reduction is due to real phenomenon or spread in the data.

Figure 2.16 (b) looks at the interactions of coolant temperature and EGR rate. It can be seen immediately that the effect of EGR is far greater than the effect of varying coolant temperature both on NO_x emissions and fuel consumption. At lower levels of EGR, the effect of coolant temperature on fuel consumption is much greater than at high EGR rates (0.4% and 0.1% respectively). Again, it is unclear from the paper whether the curvature is real or due to measurement error. Further analysis would be necessary into effects on inlet conditions and potential thermal throttling effects.



(a) *(b)*

Figure 2.16: NO_x and specific fuel consumption at 70°C and 90°C coolant temperature for (a) Injection timing and (b) EGR swings at 2670rpm, 82Nm [93]

The interactions between thermal setting, EGR rate and injection timing are important and can be exploited to find reductions in NO_x and fuel consumption through synergy effects. Up to 20% NO_x reduction was achieved with a small improvement in fuel consumption when varying all three parameters simultaneously [93].

Brace *et al.* [23] also looked at NO_x-Fuel consumption trade-offs based both on injection timing and engine thermal state but included changes in engine loading (see figure 2.17). They found that at low load, increasing the engine operating temperature[18] had a more favourable effect on NO_x/fuel consumption trade-off than shifting injection timing (compare the slopes of the temperature swings with that of injection swings). In contrast, at higher loads the effect of coolant temperature is much worse on NO_x/FC trade-off (coolant temperature swings much steeper than timing swings).

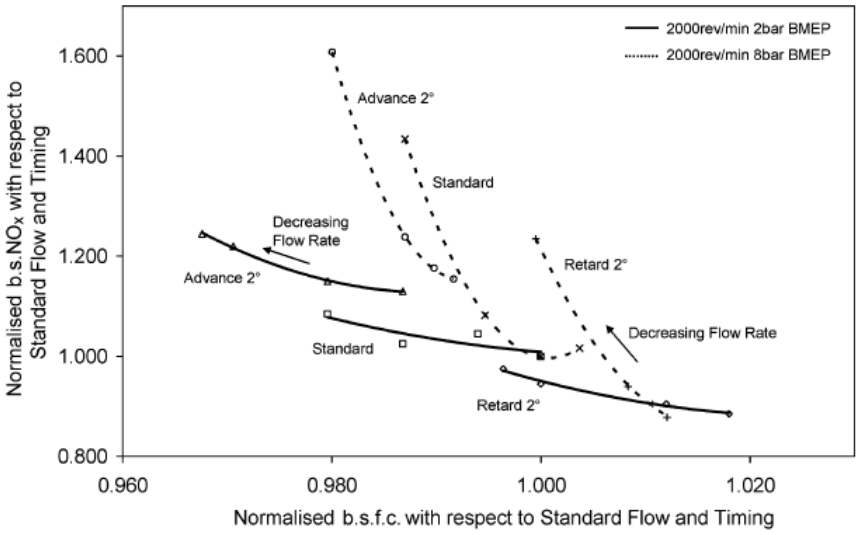


Figure 2.17: NO_x-FC trade-off based on injection timing and engine operating temperatures at low and medium loads [23] (reduced flow rates correspond to higher operating temperatures)

Brace *et al.* [18, 23] report a NO_x emissions increase of 12% at low load when coolant and oil temperatures increase by 12°C, and 33% for a similar situation at high load which are consistent with other published work [89, 93, 94]. Fuel consumption reduced by 2% and 4% in the same conditions respectively. In simple terms, it can be said the adjustments to the cooling system offer 6% and 8.25% NO_x improvement, per 1% deterioration in fuel consumption at low and high loads respectively. Injection timing offers 10% and 11.5% respectively which shows

[18] The change in engine temperature was produced by the level of throttling of the engine inlet coolant flow. Figure 2.3 showed that this approach was successful at varying steady state operating temperature.

that the thermal management system offers potential as a calibration parameter. Hydrocarbons have been found to reduce with increasing engine temperature [89], suggesting the higher operating temperature encourages more complete combustion. Experimental work by Torregrosa *et al*. [43] has also highlighted interactions with other emission species.

In-cylinder data was collected by Torregrosa *et al*. [89] to analyse IMEP with varying coolant temperature. The authors do not report significant differences in IMEP, however complete data of the ECU response to the change in coolant temperature is not published. Rate of heat release was studied for varying coolant temperatures and showed that the ignition delay was highly affected (see figure 2.18). This would seem logical as a lower wall temperature would cause more heat transfer from the cylinder and result in a lower gas temperature at the end of the compression stroke. The fuel would then take longer to reach its self-ignition temperature due to reduced heat transfer between the gas and liquid. The longer ignition delay would then cause a larger premix combustion spike. This more violent combustion would be expected to cause larger peak pressures and temperatures.

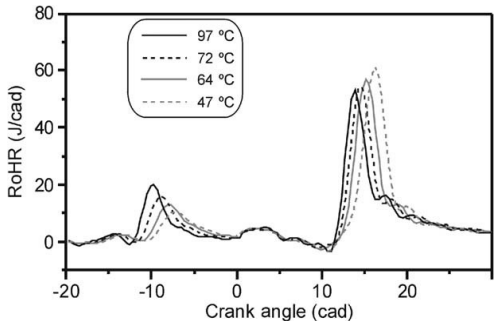


Figure 2.18: Rate of heat release for varying coolant temperatures showing increased ignition delay and higher premix spike [89]

Despite the larger premix burn, the longer ignition delay causes the process to occur later in the cycle meaning that the piston is further into its cycle as the fuel is burned. This has the effect of reducing peak cylinder pressures and temperatures. Operating the engine under cooler conditions improves the volumetric efficiency [94], meaning a larger mass of air is present in the cylinder thus increasing the heat capacity of the charge. This would reduce peak in cylinder temperatures but the effect on peak cylinder pressure is unclear as this is affected by both the temperature and the trapped mass. IMEP is therefore the result of different opposing mechanisms that seem to cancel each other out causing very little change with varying coolant temperature [89]. This study was completed at low load as no significant differences were observed at high load. This may be due to the onset of nucleate boiling

which, as discussed previously, would make metal temperatures insensitive to bulk coolant temperatures. Though no data is published with which this could be determined.

Theoretical studies of the effect of the thermal condition on the combustion events are rare, but one example is the work presented by Rakopoulos *et al*. [95]. The authors look at the effect of cylinder wall temperature on combustion model, but the setting of wall temperature is very crude and based on assumed temperature profiles. They see a small effect of wall temperature on predicted maximum in-cylinder temperature. As a single zone model was used, no accurate effect on emissions could be estimated.

5.2. Calibration potential

Emissions analysis has shown that the combustion event is affected by the engine TMS. As far as the authors are aware, there have been no comprehensive studies on the effect of coolant temperature on in-cylinder events. However, the previous work has shown the potential of the engine thermal management system to affect key performance characteristics of the engine. This highlights the potential of using the TMS as an additional input variable to the calibration process. By varying the thermal state of the engine, other parameters could be adjusted to offer gains that are not obvious when considering the TMS alone.

It is important to note that any improvements in NO_x may be traded off through other combustion control techniques to yield further benefits in fuel consumption and vice-versa. For example, a decrease in operating temperatures may allow EGR rates to be lowered which will push the NO_x emissions back towards the baseline, but at the same time yield benefits in fuel consumption and other emissions species.

Pang [25] achieved simultaneous reductions in NO_x and fuel consumption by adjusting both the thermal management and the combustion control parameters of the engine. The split cooled engine was optimised for NO_x and fuel consumption using injection timing and EGR ratio. A fully optimised engine required varying metal temperatures with operating points which would not be practical, but offered between 3% and 25% reduction in NO_x and up to 0.7% improvement in fuel consumption. A compromised solution, with constant engine temperature offered up to 14% improvement in NO_x and 0.7% in fuel consumption.

6. *Chapter summary and conclusion*

In this chapter research and production examples of engine thermal, lubricant and combustion systems and their interactions were reviewed; the following conclusions are drawn:

1) Both the lubricant and thermal systems are vital to engine operation but present a power loss which affects overall engine efficiency. These power consumptions are small relative to total engine power and benefits from improvements to these systems will be of the order of a 1% to 10%.

2) The efficiency of both lubricant and cooling systems can be improved through novel hardware which optimises the use of these systems. This is achieved through active systems that match the cooling and lubricant flow to engine requirements. These systems have reported fuel consumption gains of up to 3.5% over the NEDC.

3) Total engine friction is dominated by the piston, crank and valve train assemblies. Large discrepancies are seen depending on engine design and operating points, however friction in most engines appears to be dominated by the hydrodynamic regime. This means total engine friction is strongly dependent on oil viscosity.

4) The behaviour explained in (3) means engine friction can be reduced following cold-start if the warm-up rate is enhanced. The maximum potential can be measured by comparing engine performance under cold- and hot-start conditions over the same duty cycle; this is expected to be of the order of 4%.

5) Engine warm-up rate may be improved either through reduction of thermal inertia or by additional heat input. The reduced thermal inertia may be permanent (changes in engine design) or temporary during warm-up (fluid circuit isolation). Additional heat may be supplied from otherwise waste exhaust gases or stored from a previous hot cycle.

6) The engine thermal management system also impacts on the combustion process with a particular impact on NO_x emissions. This interaction leads to potentially unexploited benefits that can only be used if the TMS is included in the engine calibration procedure.

Chapter 3 - *Establishing methods for demonstrating small differences*

To measure small changes in fuel consumption of the order of those presented in chapter 2, a highly capable measurement process is required to demonstrate the cost-effectiveness of new hardware. Without such a system, manufacturers may be missing out on benefits that alone are small, but together could make a significant change. This chapter will lay out the experimental procedures required by this work and measure the capability of the system.

Initially the importance of repeatability and accuracy will be discussed. In section 2, the design of experiments (DoE) approach will be presented which can reduce the required experimental effort. Section 3 will discuss accuracy and repeatability issues for emissions and fuel consumption measurements. Finally, in section 4 statistical power theory will be applied as a measure of the experimental facility capabilities.

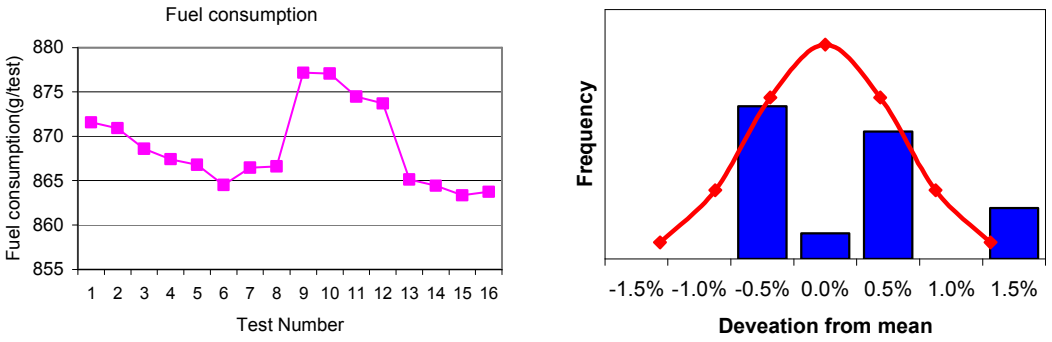
Aspects of this chapter have been published in further detail in the proceedings of the Institution of Mechanical Engineers part D: Journal of Automobile Engineering [96, 97] and been presented at an international conference [98] with references where appropriate.

1. *Workload, Accuracy and Repeatability*

1.1. Demonstrating confidence in results

When an experiment is conducted it is desirable to know if the measured value is correct. The solution to this is to repeat the experiment and compare the two results. As more and more repeat tests are performed a distribution will appear and a measure of the spread of results can be calculated. Figure 3.1 shows a typical series of fuel consumption results from a number of repeat tests. Figure 3.1 (a) shows the results for each test as they were measured while figure 3.1 (b) shows a histogram of the results with a fitted normal distribution. The spread of results can also be quantified through the standard deviation (see equation 3.1).

$$s = \sqrt{\frac{\sum (y - \bar{y})^2}{n-1}} \qquad \textbf{3.1}$$



(a) *(b)*

Figure 3.1: Line chart and distribution of fuel consumption measurements for typical testing campaign comprising 16 repeat tests

The standard deviation can be used to calculate the confidence interval, i.e. the range where the true result is expected to lie[19]. The use of confidence intervals is a good way of stating the significance of differences between two different setups. The simplest significance test is performed using Student's t distribution as described for 95% confidence interval in equation 3.2.

[19] Statistically, the recorded tests are considered a sample of a total population which represents the behaviour of the engine over its lifetime.

$$95\%C.I = \frac{s}{\sqrt{n}} \times t_{n,95\%} \qquad \textbf{3.2}$$

Where s is the standard deviation of the measurement results, n is the number of tests and t is the probability factor obtained from Student's t distribution and is a function of the desired confidence interval and the number of tests.

It can then be said that 95% of the test results should lie between the mean +/- the 95% confidence interval. When comparing two sets of data, for example from two different engine builds, the test can be used to identify statistically significant differences, i.e. to determine if observed result is a real effect, or could simply be a result of random testing variation. If small differences are to be demonstrated then the confidence intervals associated with a series of measurements also need to be small. For example, if a particular experimental procedure has 95% confidence intervals in the region of +/-1%, it will be impossible to demonstrate changes at 95% confidence that are less than 1% for a practical number of repeat tests[20]. Referring to equation 3.2, there are three ways of reducing confidence intervals:

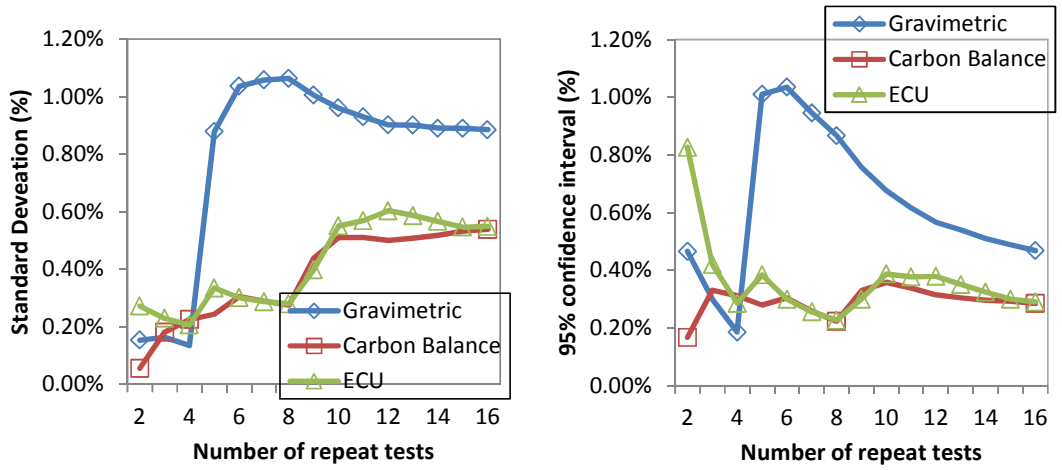
- Decreasing the standard deviation, s
- Decreasing the *t value* from Student's t-distribution
- Increasing the number of tests, n

Reducing the standard deviation of the measurement is achieved by improving test to test repeatability. Reducing this "random" variation involves identifying disturbances that influence the particular measured parameter and either inhibiting or compensating these effects. For example, in engine performance tests humidity is a limiting factor for maximum power. As this factor is hard to control and requires large investments in climatic control equipment, correction factors exist based on the measured relative humidity [99].

Decreasing the value from *Student's t-distribution* can be done in two ways: reducing the desired confidence level[21] or increasing the number of tests. The first is often a customer requirement and compromises the validity of the final result. The second is discussed below.

[20] This example has been determined using statistical power test described in later in this chapter.
[21] Typically in automotive application, a 95% confidence interval is used.

Increasing the number of repeat runs affects the equation in two ways, both directly as the denominator and indirectly through the *t value*. The effect of increasing the number of tests on the confidence interval is shown in figure 3.2: for 3 measurements of fuel consumption[22] from the testing campaign shown in figure 3.1, the standard deviation and 95% confidence interval are plotted against the number of repeat tests. As the standard deviations stabilise (figure 3.2 (a)), the confidence intervals reduce as the number of tests increases (figure 3.2 (b))[23].



(a) *(b)*

Figure 3.2:Evolution of standard deviation (a) and 95% confidence interval (b) for NEDC fuel consumption results for a typical engine testing sequence as experiments are performed

Other than increasing costs, development times and other problems including test rig availability, increasing indefinitely the number of tests causes problems of drift over time. This is clearly shown in figure 3.2 (a) by the evolution of standard deviation for the both carbon balance and ECU fuel consumption measurements. As a larger number of tests are run, the standard deviation increases as time related disturbances impact the measurement system. As a result, increasing the number of repeat runs indefinitely is not the best way of reducing confidence intervals. In addition, longer test programs are susceptible to engine failure which may render previous results useless. With cold-start testing, to ensure repeatable start temperatures facilities are usually restricted to 2-3 tests per day [100], meaning large test programs can go on for months.

[22] Details of these measurement processes will be given later in this chapter.
[23] The logic behind this is that if the number of tests was to cover the whole lifetime of the engine, then the results would be certain as data would be available for the whole lifecycle.

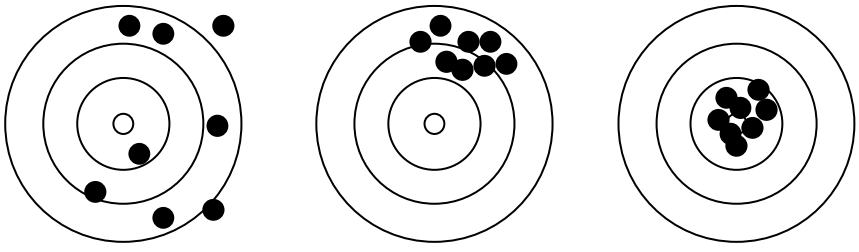
1.2. Efficient use of testing time

It is important to plan experiments such that the most information may be extracted from the least number of experimental runs. With the increasing number of parameters available on modern engines, it is neither practical nor desirable to test every possible combination and careful planning of the experiment is required. Design of experiments is a key tool that can satisfy this problem by running tests with parameters set in various combinations rather than changing one factor at a time.

1.3. Accuracy of results

Even with good repeatability and a well-planned test sequence to reduce the risk of test drift over the experimental campaign, there is no guarantee that the obtained results are correct. A very repeatable result is not necessarily an accurate result as the measurements may be consistently wrong. This is best illustrated by considering the targets in figure 3.3:

- Figure 3.3 (a) shows the performance of a poor system where the results are spread erratically around the true value.
- Figure 3.3 (b) shows a system that has good repeatability, but bias in the systems means the results is consistently wrong.
- Figure 3.3 (c) shows a system that is both repeatable and accurate. A system can be repeatable but not accurate, but a system that is not repeatable cannot be accurate.



(a) *(b)* *(c)*

Figure 3.3: Example of (a) non-repeatable, non-accurate, (b) repeatable but not accurate and (c) repeatable and accurate results with the example of a target

An experimental facility that demonstrates repeatable results without any knowledge of accuracy can be useful provided it is known that the bias or offset does not change. If results are to be compared for correlations between different facilities, then the experimental accuracy is the key link.

2. *Model based techniques*

2.1. Physical and data driven models

Model based approaches for engine development can be split into two categories: physical models and empirical, data driven models. The work in this thesis will focus on the latter, using statistical techniques based on measured data.

The trend over recent years has been to develop data driven models from steady state conditions for engine calibration and use lower fidelity physical models for engine design. For engine calibration this is expensive in testing time due to the large number of calibration variables. The Design of experiments approach has become industry standard to reduce the number of required measurements. Dynamic models are currently of interest to reduce the measurement time for each point.

Many mathematical model approaches are used where the model aim is to accurately represent the measured data, but this does not give an insight into the physical behaviour of the engine with respect to the inputs. Physical based models do give this insight, however the physical, chemical and thermal processes involved in engine combustion are too complex to be accurately calculated. There are promising results from physical modelling in some processes within the IC engine such as air path behaviour, however data driven models still prevail for emissions and fuel consumption [101].

Some novel approaches have been presented by Nebel *et al*. [102] and Shimojo *et al*. [103] which combines the physical and black box approaches (see figure 3.4). Black box dynamic models were used to predict intermediate physical values such as air fuel ratio and centre of combustion from the engine actuators[24]. The physical quantities were then used within a static black box model to predict emissions and fuel consumption. Finally the sensor dynamics were included as a final dynamic aspect to the model. The thermodynamic parameters were not disclosed in the publication, but these would be expected to include quantities such as IMEP, 50% burn, peak heat release, start of ignition etc. The data for the static model is established using a design of experiments in the design space of the thermodynamic values, meaning good control of these parameters is necessary. Dauphin *et al*. [104] have demonstrated a single

[24] This approach relies that these physical quantities can be quickly and accurately measured to allow feedback control on the experimental facility.

cylinder experimental facility capable of running the DoE in the required design space for the second phase of this modelling approach.

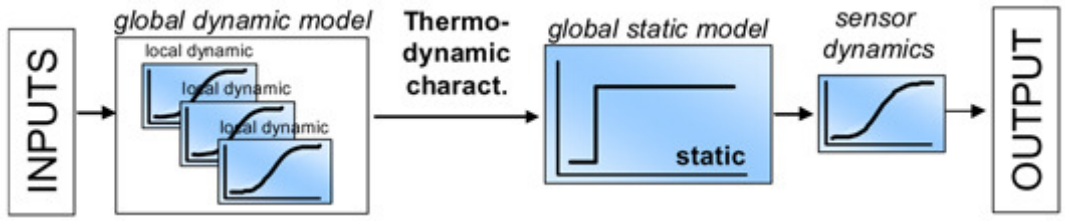


Figure 3.4: Combined physical and data driven dynamic modelling approach [102]

2.2. Design of experiments (DoE)

2.2.1. DoE process in engine development

The DoE approach aims explore a wide range of responses, both from individual parameters and multi-way interactions. To achieve this, a test plan (designed experiment) is produced to cover a region of operation (design space) by varying factors simultaneously. The simplest type of experimental design is a two level factorial design and is a good example for understanding the method. Figure 3.5 compares the DoE design space (white points) to standard one factor at a time (OFAT) approach (black points), where the three axes A, B and C represent three input variables. The main advantage of DoE is that it allows the whole design space to be explored; in contrast the OFAT method would only yield the main effects of each factor with other factors at their standard operating point (grey point). With the DoE approach, it is possible to estimate these main effects at each of the edges of the cube, giving 4 measures of these effects and an idea of the interactions between factors. If more factors are included, the number of tests can be significantly reduced if high order interactions are ignored.

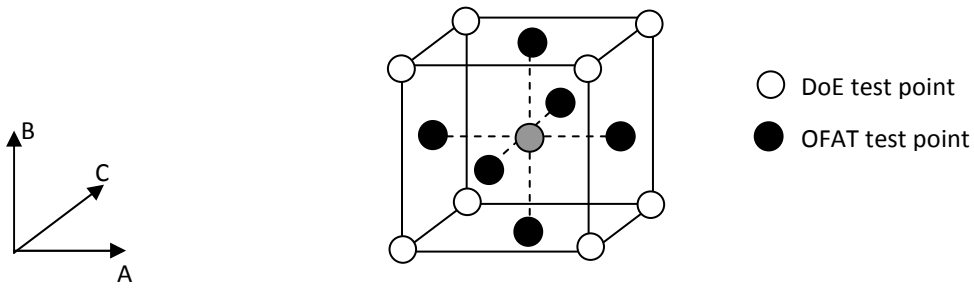


Figure 3.5: DoE and OFAT test points for 3 factor testing

Mathematically, DoE refers only to the test plan building, but in engine development it is now used to describe a complete calibration process. The method has become well established in calibration tasks [105-109] to tackle optimisation of the increasing number of systems required to meet upcoming legislations and maintain drivability and comfort [110]. Figure 3.6 summarises this DoE process from problem definition to the optimisation and verification stage.

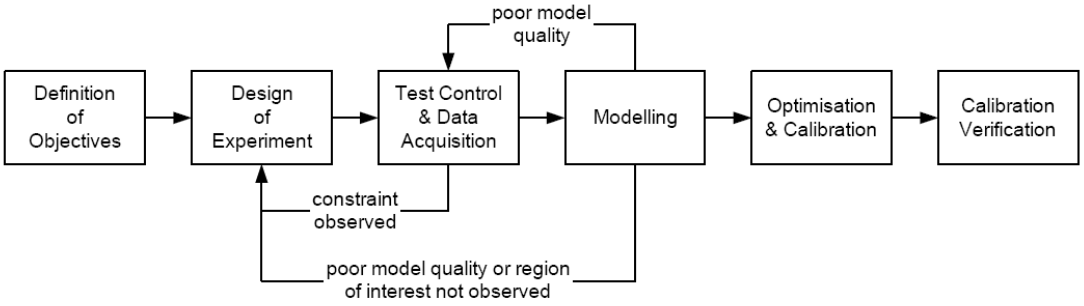


Figure 3.6: DoE model based calibration process [111]

When defining objectives it is important to consider the anticipated results as these will contribute to both test design and response model type. There are many types of designs that can be used (Full and fractional factorial designs, optimal designs, Latin hyper cubes) and a review of all designs is beyond the scope of this project for which the reader is directed to the books by Box *et al*. [112] and Atkinson and Donev [113]. The choice of design will depend on the model type, but also on experimental effort and the objectives of the task.

Polynomial models will be used extensively because of their simplicity and the scoping nature of the experiments in this work. For more detailed calibration tasks, these polynomials can encounter difficulties in representing the highly non-linear responses. During the model building process, a number of methods will be used to assist parameter selection to yield simple models with good levels of fit:

- For models with relatively small data sets, a predicted residual sum of squared (PRESS) analysis will used. This method is based on calculating a model for each subset of the data with one data point removed. A sum of squares of the error of the prediction of the removed point is then calculated. If a model parameter tends to cause over-fitting, then the PRESS value will be significantly larger as large errors will be incurred by the

removal of any single point [114]. This method is effective, but not appropriate for large data sets because of the computing effort required.

- Stepwise parameter selection: this is based on assessing the confidence interval of regression coefficients. The method can be applied backwards , starting with a full model with all coefficients and iteratively removing the least significant term until all terms are significant, or forwards starting with no coefficients and adding most significant terms until no more significant terms are available [115].
- For models with large data sets and many degrees of freedom, it is impractical to use the PRESS or stepwise methods due to large computing times required and the orthogonal least squares (OLS) approach will be taken. This parameter selection technique is based on assessing the correlations between output and individual columns of the regression matrix meaning actual regression is not required for parameter selection.

Various statistical techniques are available for assessing the goodness of fit and the predictive power of the model:

- Coefficient of determination (R^2), which is the ratio of variation explained by the model to the total variation in the data. This can only be used to measure goodness of fit as an increase in number of model parameters will always lead to increased explanation for variation by the model.
- Adjusted R^2, which adjusts the R^2 value to account for the number of degrees of freedom, hence a model with more terms will only increase the adjusted R^2 value if significantly more variation is explained.
- Residuals analysis [116]
- PRESS analysis [117]
- Model validation using independent data set

With validated models, optimisation routines can be used to assess the best calibrations/setups with the modelled system taking into account output constraints and the valid training regions of the model[25].

[25] For experimental designs with a large number of input variables the determination of the valid input space should be assessed using computational geometry and the calculation of a convex hull. This will be illustrated in chapter 9.

2.2.2. Extension of DoE to dynamic engine modelling

Dynamic design of experiments follows the same principles but extends to the frequency response as well as the steady state. The aim is to build models that are accurate during simulation of transient events for optimisation of engine calibration over these periods. The European homologation cycle is relatively steady state and only two transients are really problematic for NO_x calibration. On the other hand, the American and Japanese homologation tests are much more transient in nature [5] and require much more attention to dynamic events. It is interesting from a simulation perspective to develop empirical models that are not duty cycle specific and valid over a range of homologation cycles.

From a modelling perspective, rather than a sequence of steady state points, the test points become a dynamic time series and the model takes a particular structure to account for the dynamic aspect of the response. The procedure is similar to that of system identification in control theory. Because the dynamic models are validated using dynamic measurements, the transient responses of the various sensors also need to be accounted for within the model. This represents the main gains in experimental time as the settling time for each measurement point is no longer required: design of experiments reduces testing time by reducing the number of test points, dynamic measurements can reduce the time of the measurements themselves, further reducing development time.

(a) Dynamic models

Non-Linear Auto-Regressive Moving Average models with Exogenous inputs (NARMAX models) is the general designation for dynamic empirical models that incorporate memory effects whereby the behaviour of the system depends not only on the current state of inputs but also the previous states. The model output can also be affected by previous output values which are the main factors in modelling sensor dynamics.

Guhmann and Riedel [118] compared a number of different modelling approaches on a single training and data set. The types of models investigated included neural networks, polynomial approaches such as parametric Volterra series and Radial Base Function (RBF) networks. The models were used to simulate NO_x and HC emissions. The performance of each was very similar, although the Volterra series gave the smallest RMSE for NO_x emissions prediction. For HC emissions a *neural network using hyperbolic tangent sigmoid activation functions with output feedback* performed best, although the parametric Volterra series gave similar

behaviour and ranked 2nd out of the 10 models for prediction RMSE. In this study both the training data and the validation data were issued from the same test, with 2/3 of the data used for training and the rest for validation. The test data was recorded using sinusoidal excitation signals and consequently both the training and validation data were of the same type. In real calibration tasks, the models would be expected to perform well over the homologation drive cycles and therefore the performance of the models may be different over other validation cycles with significantly different shapes[26].

Because the work in this thesis is focused on emissions modelling, a polynomial approach was taken using the Volterra series. In addition to the work by Guhmann and Riedel [118], other authors have successfully modelled engine emissions using this approach [119]. The Volterra series is described by equation 3.3 and illustrated in figure 3.7.

$$\hat{y}(k) = \begin{pmatrix} f([x_1(k), x_2(k), \ldots, x_n(k)]) \\ + f([x_1(k-1), x_2(k-1), \ldots, x_n(k-1)], [x_1(k-2), x_2(k-2), \ldots, x_n(k-2)], \ldots) \\ + f(y(k-1)) \end{pmatrix} \quad \textbf{3.3}$$

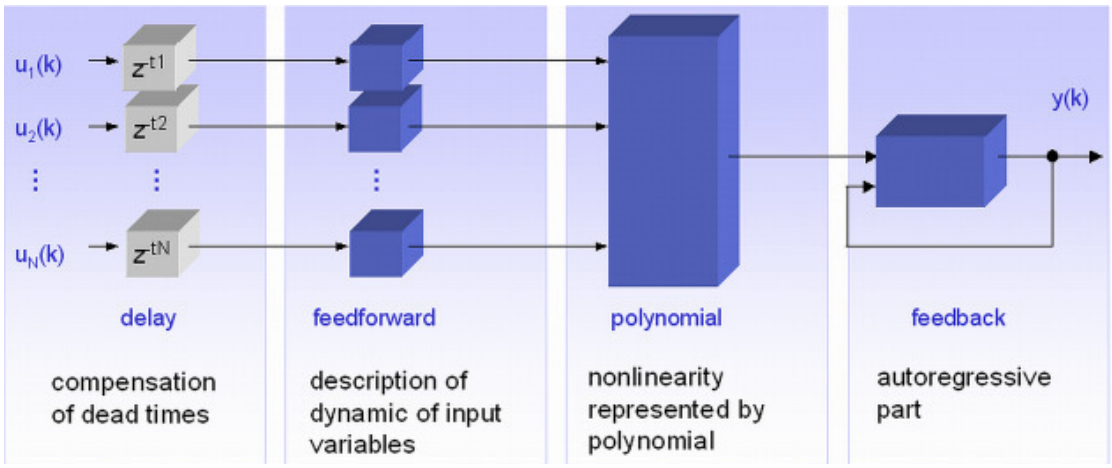


Figure 3.7: Graphical representation of the Volterra series for dynamic modelling [120]

Because the Volterra series is polynomial, there are no complex training sequences and standard least squares can be used with some form of parameter selection to estimate coefficients for the first two lines in equation 3.3. The only issue is the many parameters and

[26] The NEDC is much more steady state and consequently very different to the sinusoidal excitations. The FTP and Japanese cycles are more sinusoidal in nature and may be more representative of this validation, so long as the frequency range is appropriate.

large data-set mean some useful parameter selection methods like PRESS analysis are not suitable because of computing times. Orthogonal least squares performed adequately and was used for this work. The third line accounts for the sensor dynamics and represents the autoregressive aspect of the model. Because of the strong correlation between the output y_k and previous output y_{k-1}, least squares is not suitable and an optimisation routine is required to estimate the final coefficient. Without this autoregressive aspect, the model would account for sensor response times by underestimating the actual engine response. This is best illustrated by considering the response to a square wave shown in figure 3.8:

- In figure 3.8 (a), a measurement response is shown (red) following a step change in input (black). Before the system can settle, the input is returned to its original state. If the sensor dynamics are not modelled, then the static response model has to account for the dynamics by underestimating the final value. This is shown in blue and is effectively the average value over the sample period.
- Figure 3.8 (b) shows the same response, but the static model is adjusted because the sensor dynamics are accounted for by the autoregressive part of the model. Each of the regression coefficients need to be adjusted and this is treated as an optimisation problem where the model structure remains fixed but the coefficients are allowed to float while the fit RMSE is minimised.

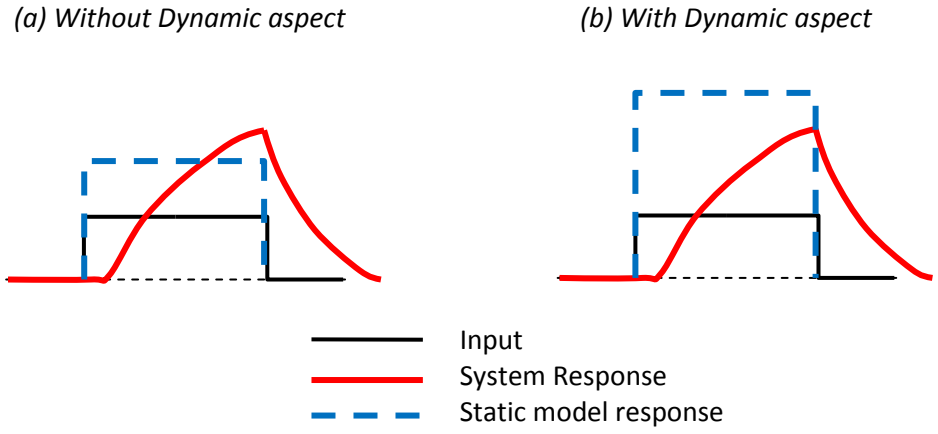


Figure 3.8: Illustration of static model prediction (a) without and (b) with the inclusion of sensor dynamics modelling for a step input.

(b) Model Training signals

Two excitation signals are most common for recording training data for dynamic engine modelling: *amplitude modulated pseudo random binary signals* (APRBS) [103, 121, 122] and *swept frequency*[27] *sinusoidal signals* (Chirp) [119]. The APRBS is composed of a series of step changes with varying hold times; an example for engine speed is shown in figure 3.9.

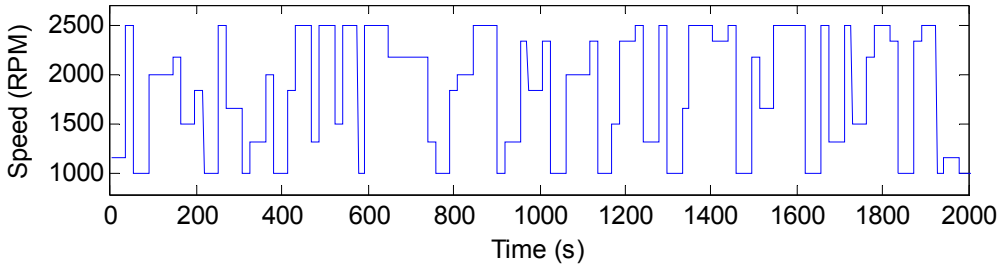


Figure 3.9: Example APRBS signal for engine speed covering NEDC design space

There are practical issues for implementing the APRBS as this sequence involves harsh changes in operating points which are not representative of real engine use. In addition, these operating conditions can be dangerous for the engine and experimental facility. Some examples are illustrated in figure 3.10 which shows demand and measured engine speed and torque for a trial APRBS. The requested speed signal followed well, however there are a number of issues relating to the torque signal in response to certain signal requests:

1. **Step down in engine speed**: when the host system requests a drop in speed, it can deploy all available brake torque against engine. This results in a spike in drive shaft torque that will be inversely proportional to the ramp time and can be dangerous for the engine and testing hardware. The signal can be developed to include ramp times between operating points, however this is not straight forward as this will affect the engine response.
2. **Step up in torque demand**: the response from the test control system is to apply maximum pedal position to achieve this torque as soon as possible. This often results in an overshoot of the desired value. This situation is not dangerous for the hardware as the maximum torque is limited by the engine fuelling, however from a testing perspective it means that all increases in torque will result in near-to maximum fuelling, regardless of the target step.

[27] The frequency variation is controlled by a simple time based mathematical function. This function is typically linear or logarithmic; figure 3.11 shows a Chirp signal obtained from a logarithmic frequency variation over 1000 seconds. This complete signal is obtained by mirroring this base portion.

Publications showing engine modelling using these excitations sequences rarely mention these difficulties. Shimojo *et al*. [103] may have reduced these effects by performing local experiments at different point in the speed and load map.

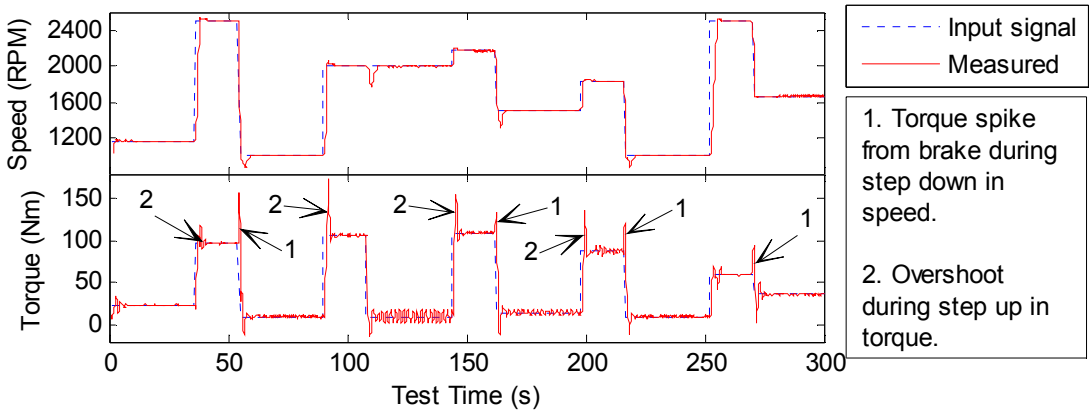


Figure 3.10: Torque and pedal behaviour when running torque based APRBS

An example chirp signal is shown in figure 3.11: this continuous signal has the advantage of being gentler on the engine operation and closer to real driving situations.

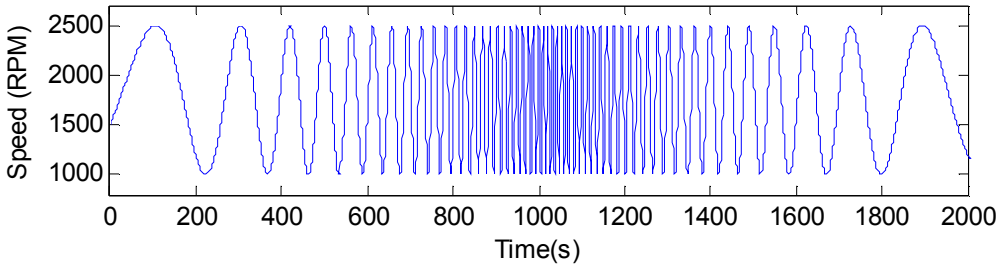
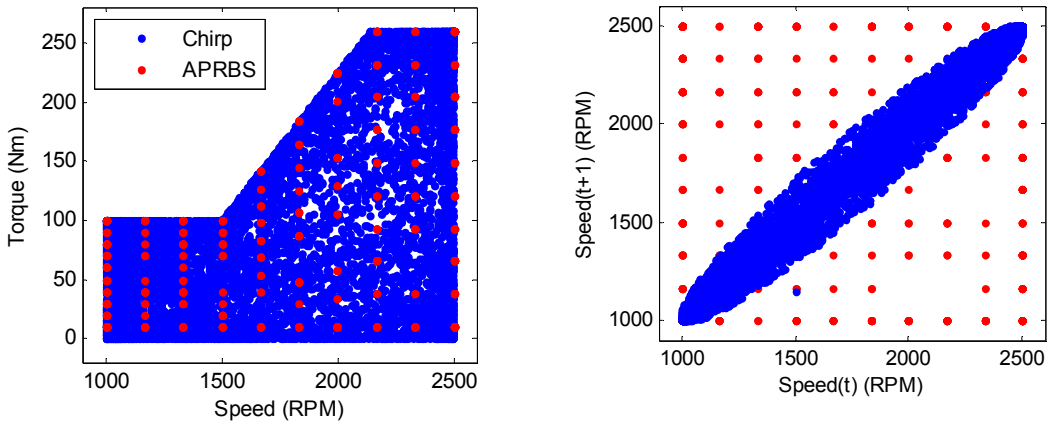


Figure 3.11: Example Chirp signal for engine speed covering NEDC design space

As with the standard design of experiments, it is important to use the model structure to design the experiment. It is desirable to have a test plan that will give little correlation between the inputs, but also between the current and previous time steps for each input values. Figure 3.12 illustrates these correlations for both APRBS and chirp signals:

- Figure 3.12 (a) shows the 2D scatter plot for torque and engine speed for signals aimed to replicate the NEDC operating space.
- Figure 3.12 (b) shows the correlation between current and previous time steps for engine speed.



(a) *(b)*

Figure 3.12: Comparison of APRBS and Chirp signal design space coverage for (a) two independent variables and (b) current and previous time steps for one input variables

Both signals perform well for design space coverage between the two input variables. The APRBS has very good design space coverage for current and previous time steps which probably explains why it is often used in system identification tasks[28]. There is a correlation for the Chirp input and the plotted shape is worth discussing:

- If a zero frequency signal was used (i.e. a ramp input), there would be a perfect correlation.
- If a single frequency sinusoid was used, an ellipse would be produced.
- The varying frequency signal lies somewhere between these two, with the coverage dependent on the range of frequencies.

Despite the strong correlation, authors [119] have successfully produced Volterra models, provided this covers the operating range. Baumann *et al*. [123] have compared these two and other excitation signals in simulation and concluded that sinusoidal based signals performed better in terms of final model quality in the context of automotive engines. For these reasons, this work will exclusively use the sinusoidal based signals.

While the approach has been effective for modelling engine response to common calibration variables that are easily controllable via the ECU or host system, the difficulty arises in capturing the behaviour of the engine under dynamic thermal events such as cold-start. Unlike engine speed, load and ECU controls, the engine temperature cannot be varied freely

[28] Many systems are less sensitive to this input type. If engine speed and torque can be omitted from the required inputs then this signal may be considered for use in IC engines.

according to an input signal if realistic temperature distributions are to be reproduced. Using in-cell engine cooling devices it is possible to force the engine coolant temperature to a wide range of temperatures, however it is important to note the differences between an engine during warm-up and a hot engine with artificially low coolant temperature. The work in this thesis will aim to replicate some of the engine identification work under fully warm conditions, but build on this by applying a simple method to capture the dynamic thermal behaviour of the engine.

3. *Improved measurement systems*

3.1. Accuracy of fuel consumption measurement

3.1.1. Fuel Consumption measurement principles

A number of well-established measurement methods are available commercially for direct or indirect measurement of fuel consumption [124] based on gravimetric or volumetric principles, equating carbon content of exhaust gases or through ECU data. Because of equipment costs it is rare for facilities to have more than one direct measurement of fuel consumption, however most facilities will have emissions analysers and they offer a good backup to the primary measurement. At the test facility used in this study, a gravimetric fuel balance and emissions analysers were available:

- **Gravimetric fuel balance**: the mass of fuel flowing to the engine is measured directly on a cumulative basis. The weight of fuel in a supply beaker is continuously measured as fuel flows from the beaker to the engine. The resultant change in weight is the measured fuel consumption. Because the beaker is of fixed capacity, refilling is required and measurement must be suspended.
- **Carbon Balance**: an indirect estimate of fuel consumption can be obtained by equating the mass of carbon in the exhaust to the carbon concentration of the fuel. Carbon in the exhaust is estimated from CO_2, CO and HC measurements. The carbon content of the fuel is obtained from the fuel supplier or measured using chemical analysis. The emissions are sampled on a continuous basis just before the catalyst and after correct time alignment [125, 126] to give details of transient events. Whilst this remains an indirect estimate, it serves as a good backup for the primary measurement.

Results from five experimental campaigns at this facility form the basis for this work. For each setup, multiple tests were run over the NEDC either from cold-start, following an overnight soak at 25°C, or hot-start, following a warm-up procedure. It is important to note that regardless the measurement technique, the *true and unknown* fuel consumption is the same and the measured estimate will be the result of both random disturbances and bias in the methods. Random effects will be identified and minimised by multiple test runs, leaving the remaining measurement system bias.

Figure 3.13 shows typical cumulative fuel consumption measurements over an NEDC for both a cold- and hot-start test. Cold-start fuel consumption is expected to be higher as a result of lower oil temperatures which cause higher engine friction (see chapter 2). As the engine warms up the cold-start fuel consumption is expected to drop to the same level as that of the hot starting engine. This is the case for carbon balance fuel consumption measurement where the difference in cumulative fuel consumption stabilises to 40g. Result from the gravimetric measurement follow a similar trend over the first 600 seconds, however over the next 600 seconds the results suggest fuel consumption is lower in the cold-start test. Nothing in engine operation is typical of this behaviour, suggesting a disturbance to the gravimetric measurement.

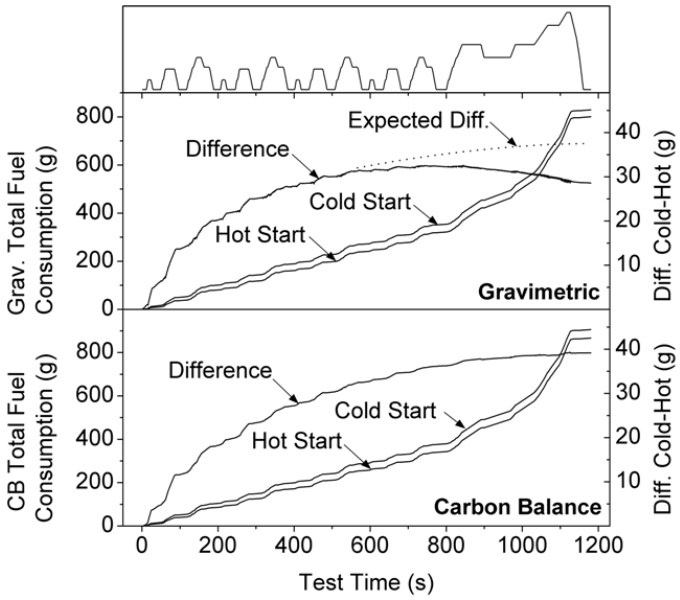


Figure 3.13: Gravimetric and carbon balance fuel consumption over a hot and cold-start drive cycle, including difference between cold- and hot-start tests

3.1.2. Detailed analysis of carbon balance method

Overall fuel consumption is obtained using equation 3.4, where *0.428* and *0.273* represent the ratio of atomic weight of carbon to the molecular weight of carbon monoxide and carbon dioxide respectively.

$$FC_{CB} = \frac{1}{\frac{w_C}{100}}\left(\frac{w_C}{100} \times m_{THC} + 0.428 \times m_{CO} + 0.273 \times m_{CO_2}\right) \quad \textbf{3.4}$$

From the terms in equation 3.4, the carbon balance estimate is dependent on knowledge of the fuel properties and the mass emissions of HC, CO and CO_2. This highlights the need to use reference fuels and accurate measurements of exhaust emissions by mass. Emissions analysers measure the concentrations of their respective species by volume and post processing yields an estimate by mass (equation 3.5). The exhaust mass flow rate must be estimated in addition to the ratio of densities of exhaust species to that of the total exhaust (see BSi standards [127]).

$$\dot{m}_X = \dot{m}_{ex} \times c_X \times \rho_{Ratio,X} = \dot{m}_{ex} \times \frac{v_X}{v_{ex}} \times \frac{\rho_X}{\rho_{ex}} \times K \quad \textbf{3.5}$$

It is important to note the conditions of temperature and humidity where each of the terms in equation 2:

- **Exhaust mass flow**: estimated as the sum of measured intake air flow and fuel flow, therefore includes all water vapour both from ambient air and combustion
- **Emissions concentrations**: emissions analysers measure under different conditions depending on emissions species. CO and CO_2 are measured by non-dispersive infra-red spectroscopy (NDIR) which analyses the light spectrum after certain frequencies have been absorbed by the gas. As the absorption spectrum of water vapour interferes with that of CO_2 and CO, the sample exhaust gases are cooled to remove the water vapour by condensation. In contrast, HC emissions are measured by flame ionisation (FID) [124] and are not affected by the presence of water and are measured under hot and humid conditions.
- **Relative density**: this term is defined under standard conditions from molecular mass and volume at 0^oC and 0% humidity [127].

BSi suggests adjusting the measured exhaust species concentration to account for the lack of water vapour [127]. Another correction factor has also been developed by the author of this thesis [97], however overall results were similar. This correction factor *K* reduces the measured volumetric concentration slightly to account for the volume of water vapour in the exhaust flow (equations 3.6 and 3.7).

$$K = \frac{c_{X,wet}}{c_{X,dry}} = \frac{\dot{v}_{ex,dry}}{\dot{v}_{ex,wet}} = 1 - \frac{\dot{v}_{H_2O,cond}}{\dot{v}_{ex,wet}}$$

$$= 1.008 \times \left(1 - \frac{1.2442 \times H_a + 111.19 \times w_H \times \frac{\dot{m}_f}{\dot{m}_{air,dry}}}{773.4 + 1.2442 \times H_a + \frac{\dot{m}_f}{\dot{m}_{air,dry}} \times f_{fw} \times 1000} \right) \quad \textbf{3.6}$$

where

$$f_{fw} = 0.055594 \times w_H + 0.0080021 \times w_N + 0.0070046 \times w_O \quad \textbf{3.7}$$

$v_{H2O,cond}$ *is the volume of water vapour that condenses in the analyser cooler.*

f_{fw} *is the volume change from combustion air to wet exhaust air per kg of fuel.*

The correction factor *K* is presented as the ratio of emissions species concentration in wet exhaust to the concentration in dry exhaust. This is equivalent to the ratio of dry exhaust to wet exhaust, on a volumetric basis. This accounts for the ambient water vapour and water from combustion that condensate in the analyser cooler.

3.1.3. Detailed analysis of Gravimetric method

In contrast to the carbon balance approach, this fuel consumption estimate is based on a single measurement. The fuel consumption between any two instances in time is the change of fuel weight in the beaker over that time period (see equation 3.8). The gravimetric fuel balance itself has a high accuracy (±0.05% or ±0.03g) and a built in calibration procedure using a calibrated weight to ensure the device performs well over time. Inaccuracies result from the installation of the device within the experimental system.

$$FC_{grav} = m_{f,bea,mes,start} - m_{f,bea,mes,end} \quad \textbf{3.8}$$

Figure 3.14 shows the fuel supply circuit from the measurement balance with the evolution of fuel temperatures over the NEDC: fuel is supplied from the fuel balance to the engine using a gravity feed. It should be noted that the instrumentation measures fuel flow into the supply circuit, not into the combustion chamber. After filtration, fuel is pumped to a high pressure in the common rail which subsequently supplies the injectors. Such is the design of the electro-hydraulic injectors[29] combined with leakage from the pump and rail, a significant amount of fuel is returned via a low pressure *spill circuit.* As suggested by Stone [52], this is returned downstream of the fuel beaker, however other authors [124] give contradicting recommendations. By the same process the fuel acts as a cooling medium for the injection system and a fuel cooler is required to avoid excessively high fuel temperatures.

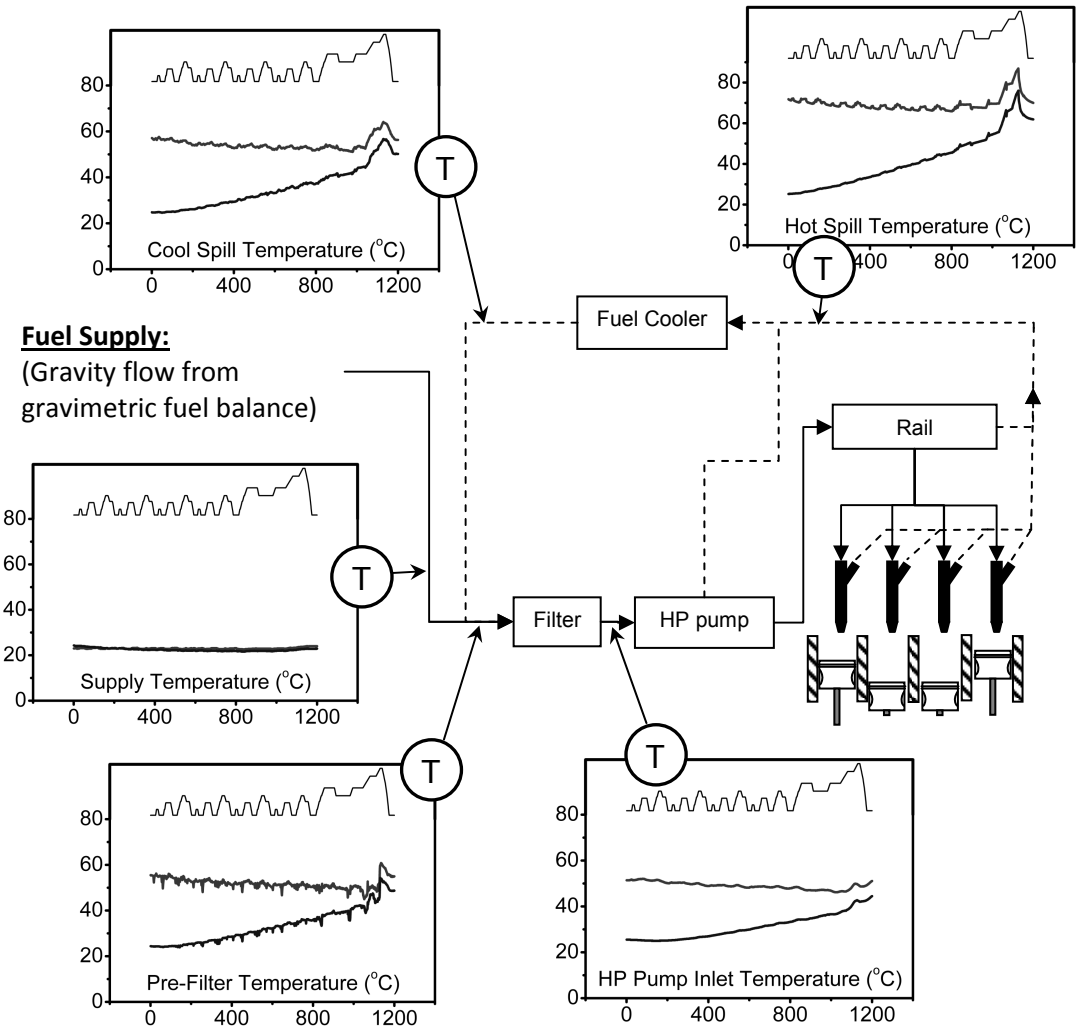


Figure 3.14: Fuel temperatures in the fuel supply circuit for hot- and cold-start tests (hot-start tests represent the higher, constant temperatures)

[29] Typical injector design for EURO IV engines used a return flow of high pressure fuel to act as a hydraulic amplifier for a solenoid actuator to execute needle lift (see Guerrassi and Dupraz [128]).

Significant temperature gradients exist during the cold-start test with temperatures in the spill circuit before and after the cooler rising by approximately 35°C and 25°C respectively. This rise in temperatures will reduce the density, and hence the total mass of the fuel in the circuit. If the mass of fuel downstream of the fuel beaker changes then the fuel consumption measured by the fuel balance will be wrong. It must be emphasised that the flow meter measures fuel flow into the fuelling circuit, not necessarily fuel consumed by the engine.

To illustrate this effect, one should consider the situation where fuel is not flowing: an increase in temperature would cause fuel expansion, pushing fuel back into the beaker. This would result in a measurement of negative fuel consumption, despite no real fuel use. When fuel is flowing, this does not result in a negative reading, but an underestimate of true fuel consumption. On the other hand, thermal expansion will also increase pipe volumes which will have the opposite effect on fuel mass in the circuit. Finally significant pressure drops may cause air release, resulting in further fuel expansion.

Thermal expansion may be corrected in post processing by calculating the change in fuel mass downstream of the beaker. The fuel circuit was split into known volumes [97] and a representative temperature was measured for each (figure 3.14). Because of the high pressures in the fuelling circuit, the correction factor was also designed to incorporate this effect[30].

The temperature/pressure-density characteristics of the fuel in this study has been derived from the work published by Dzida and Prusakiewicz [129] and Rodriguez-Anton *et al*. [130] and is summarised in figure 3.15. Although some extrapolation was required for completeness, this was not extensively used for the NEDC. For further details relating to this data the reader is directed to the authors publication [97].

[30] Although 5 fuel temperature measurements were available, only a single pressure measurement was take in the high pressure rail. Each portion of the circuit was considered to either be at the high measured *rail* pressure or a low *supply* pressure, calculated based on the height of the fuel balance and the fuel density.

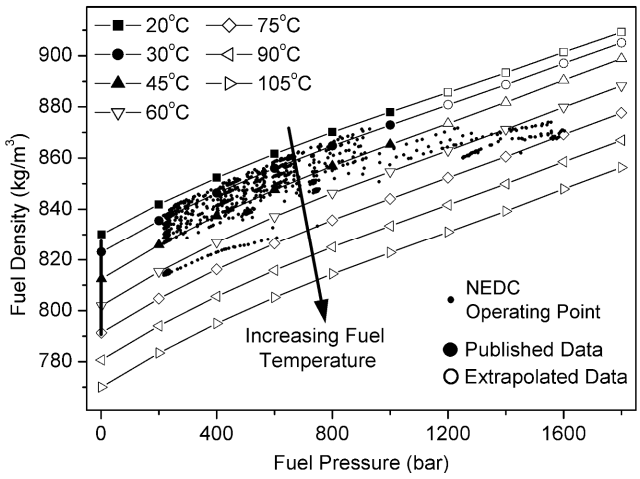


Figure 3.15: Fuel density with respect to pressure and temperature, showing published data [129] and extrapolated region. All NEDC operating points are superimposed to show small reliance on extrapolated data

The mass of fuel in the circuit was determined by summing up the calculated masses of fuel in each of the five volumes. This was then compared to the mass of fuel in the circuit at the beginning of the test and used to correct the measured beaker mass (equations 3.9 and 3.10). The corrected beaker mass was then substituted into equation 3.8.

$$\Delta m_{f,c} = m_{f,cir,t=0} - m_{f,cir,t=t} \qquad \textbf{3.9}$$

$$m_{f,bea,corr} = m_{f,bea,mes} - \Delta m_{f,cir} \qquad \textbf{3.10}$$

Further analysis was conducted to estimate pipe expansion[31] which clearly also affects the circuit volume and an air trap was installed to assess any bubbles that may form as the fuel undergoes the large pressure drops in the injector, rail pressure relief valve and leaking from the pump. The effect of pipe expansion was found to be 9% that of fuel expansion but no significant volume of air was collected in the air trap [97].

For the same set of cold and hot NEDC tests in figure 3.13, the effects of gravimetric fuel consumption correction factor are presented in figure 3.16. The shape of the curve representing the difference between hot- and cold-start tests is now consistent with the carbon balance method. The majority of the 15g increase was a result of changes in fuel temperature changes and the effect of pressure was negligible. The hot-start test was not affected by the correction factor as no significant net change in temperature occurs over the cycle.

[31] Pipe expansion was estimated based on the longitudinal and radial expansion of the rubber hoses assuming these were at the same temperature as the fuel. For full details of the calculation the reader is directed to the author's full publication on fuel consumption accuracy [97].

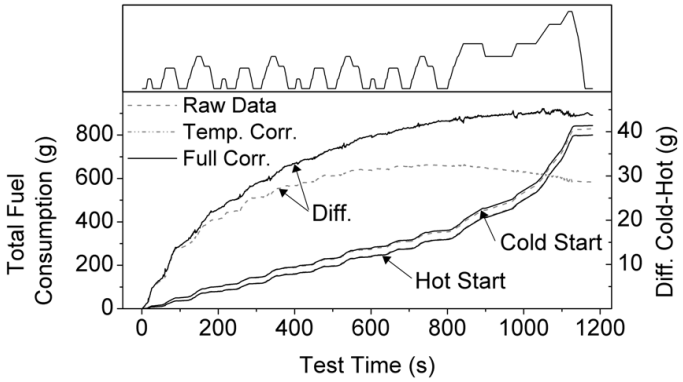


Figure 3.16: Raw, temperature-corrected and fully- corrected (temperature and pressure) fuel consumption for hot and cold-start NEDC

It was surprising that the effect of pressure was insignificant compared to that of temperature. On closer inspection, the proportion of the fuel circuit at pressure represents only 5% of the total volume and the average rail pressure during the NEDC is only 550bar. Whilst insignificant under these conditions, on a different engine, working at higher pressures with a larger high pressure fuel volume and operating nearer to full load for extended periods, this effect may need to be considered. For a particular NEDC cycle, the fuelling circuit was allowed to cool for 4 hours after the test[32] and the changes in apparent fuel consumption were measured. Over this period the measured beaker mass reduced by 30g, clearly highlighting the thermal effects on the process [97].

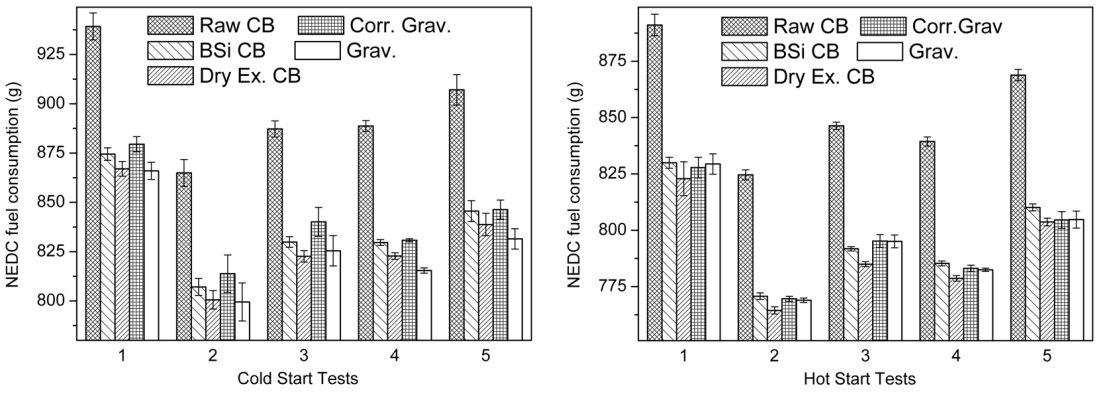
3.1.4. Summary of accuracy benefits

With the exception of the post-test cool down, the various methods have been applied retrospectively to the five experimental campaigns (see figure 3.17). The raw measures of carbon balance and gravimetric fuel consumption showed an offset of 50-80g (7-10%) over the NEDC, however the application of various correction factors, derived from first principles, has improved agreement between the methods.

Omitting the correction factor in the emissions mass results causes an overestimate in fuel consumption of up to 8%. The correction factor also seemed to improve the test to test repeatability as confidence intervals were smaller for corrected results: this is thought to be the result of including ambient humidity in the measurement process.

[32] The engine was allowed to cool back to its start temperature whilst the fuel temperatures and gravimetric fuel balance measurements were recorded. During this time fluctuations in fuel temperature as a result of the cell air conditioning system were also recorded on the gravimetric fuel balance highlighting the strong thermal link with the measurement process.

The gravimetric correction increases the raw measure by 10-15g for cold-start tests, but does not significantly change the hot-start tests. For the experimental setups considered here, best agreement appears to be between the BSi corrected carbon balance and the corrected gravimetric methods. The difference between these two methods is less than 10g (1%) and in most cases there the difference is not statistically significant.



(a) *(b)*

Figure 3.17: Fuel consumption estimates for five experimental setups and for (a) cold and (b) hot-start tests for raw carbon balance (Raw CB), British Standard corrected carbon balance (BSi CB), dry exhaust corrected carbon balance (Dry Ex. CB), Temperature and pressure corrected gravimetric (Corr. Grav.) and raw gravimetric (Raw Grav.)

With large discrepancies between measurements, the test operator is faced with deciding which method is most trusted. In contrast, with significantly better agreement and no statistical differences, confidence is increased in both methods *and* the final result. Finally, with better agreement, these offer a good platform to calibrate ECU fuel consumption estimate for accurate and detailed analysis transient of high speed transient events.

3.2. Experimental repeatability

3.2.1. The contribution of ambient conditions

The installation of the engine on a dynamometer removes some sources of variability that can occur when an engine is tested on vehicle. A significant number of components are removed from the system such as gear boxes, final drives, tyres and the driver that have an negative impact on the repeatability of the experiments [96, 98]. Control can be automated allowing for very repeatable engine operation. Because the engine is operated within a building, variations in ambient conditions can be reduced. More advanced facilities will have high performance temperature control because of the effects of engine start temperature on both the physical

behaviour (friction) and the engine control system. Through the same control system, the effects of intake air temperature on engine behaviour are also accounted for. However, in the system at the University of Bath, as many other facilities, there is no control for humidity. The ambient air in the cell is also used for combustion and is subject to the current weather conditions.

The effect of ambient conditions on both NO_x emissions and engine performance has been known for some time. To achieve comparable testing results despite day to day and location to location variations, correction factors have been developed to adjust results to standardised conditions [131].

NO_x emissions correction factors for ambient temperature and humidity have been established [132] and undergone reviews [131, 133] and updates [127] for both Gasoline and Diesel engines. Chapter 2 briefly presented the NO_x formation process and the links with in-cylinder gas temperatures. Some authors have demonstrated a strong exponential relationship between thermal NO formation to the adiabatic flame temperature[33] [135]. From a theoretical perspective if dissociation reactions are neglected, ambient water vapour is an inert gas and reduces the adiabatic flame temperature. If the endothermic dissociation reactions are also considered, then there will be a further reduction in flame temperature. The additional water vapour also reduces mixing between the fuel and oxygen. The effect of this is higher NO_x production in dryer conditions and lower NO_x production in humid conditions. A well-known correction factor from the literature is presented in equation 3.11 [132][34]

It is important to note the difference between this correction factor and that presented previously for dry/wet concentrations of CO and CO_2. The previous correction factor is applied to represent the engine behaviour during that particular test. The NO_x correction factor effectively applies an offset from the actual measured behaviour during this particular test to

[33] Adiabatic flame temperature is a theoretical temperature calculated from the reactants and products of combustion assuming complete combustion, and assuming no heat loss, i.e. all heat from combustion results in an increase in temperature of the products [134]. Other authors have shown that under different conditions other factors relating to local temperatures, oxygen concentrations and flame size also affect thermal NO formation [135].

[34] The correction factor was superseded over the duration of this project by a simpler relationship, however to preserve consistency throughout the project this was not implemented. A comparison of the two methods was conducted at the end of the projected on a selected number of experiments and shows little difference between the two for the ranges of ambient conditions experienced over the course of this work. $K_{NOx} = \frac{1}{1 - 0.0182 \times (H_a - 10.71) + 0.0045 \times (T_a - 298)}$

predict the performance of the engine had the test been conducted under standard conditions (ambient temperature 25°C and absolute humidity 10.71g/kg dry air).

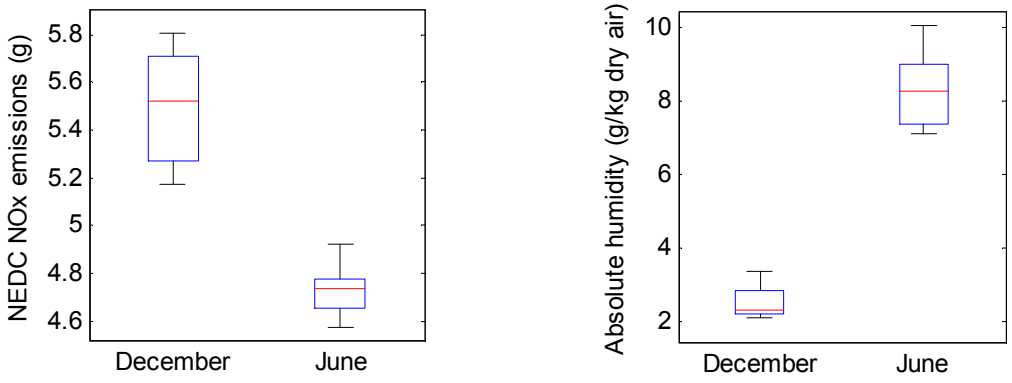
$$K_{NOx} = \frac{1}{1 + A \times (H_a - 10.71) + B \times (T_a - 298)} \quad \textbf{3.11}$$

where

$$A = 0.309 \dot{m}_f / \dot{m}_{AIR,dry} - 0.0266$$

$$B = -0.209 \dot{m}_f / \dot{m}_{AIR,dry} + 0.00954$$

Cold-start experimental campaigns will often spread over many months because of the thermal soak times required. Consequently, in the UK and most parts of the world testing campaigns are subject to significant changes in ambient conditions between cold winter months and warm summer. The large variations in outside temperature cause large variations in absolute humidity which result in large changes even in temperature controlled environments. The testing campaign presented in this thesis has spanned approximately a 9 month period; opportunistically, a particular engine build has been tested both in December and June with a number of repeat NEDC. NO_x results from these tests are shown in figure 3.18 along with the absolute humidity level. In both cases the NO_x correction factor described in equation 3.11 has been applied.



(a) *(b)*

Figure 3.18: Comparison of (a) NO_x emissions and (b) humidity for repeated engine tests under the same hardware configurations in December and June.

Clearly there are large discrepancies between results despite the use of the correction factor. It would be naïve to conclude that all the variation in NO_x emissions was a result of the variations in humidity levels, however close inspection of many other measured parameters such as injection timing, EGR rates, and temperatures has showed little difference between

the two campaigns. This does not present conclusive evidence but does suggest that the NO_x correction factor is not aggressive enough for the current engine.

Using the results from this data it is possible to estimate an empirical NO_x correction factor for the testing in this particular project by assuming a linear relationship between humidity and NO_x. Such a relationship is illustrated in figure 3.19 and described by equation 3.12. This relationship is obtained by regression of a linear relationship to the data and statistical analysis shows that the relationship is statistically significant. The standard conditions for this equation were chosen to be 2.5g/kg dry air. This is significantly different to those from the correction factor described above, but represents the conditions at which the main experimental work was performed for this thesis. In this way the correction factor is chosen to affect the smallest number of tests. Clearly this approach is specific to this engine and this project. Also, because this relationship remains untested, it is important to refer to the raw measured values in conjunction with the corrected values.

$$NOx_{corr} = NOx \times \left(1 - 0.0218 \times H_A\right) \qquad \textbf{3.12}$$

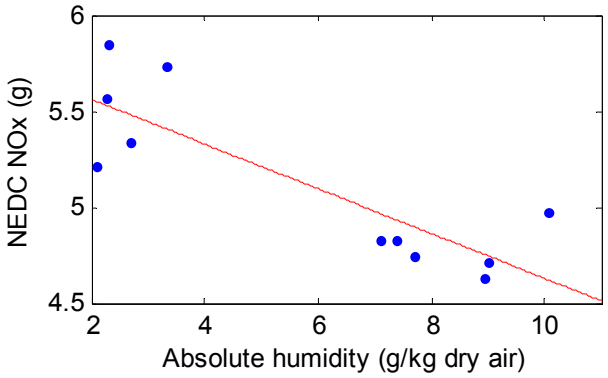


Figure 3.19: Linear fit for NO_x and humidity for determining empirical NO_x correction factor

Engine power correction factors exist to account for changes in ambient temperature, pressure and humidity [52, 136]. One common use for these correction factors is to determine the maximum engine power which represents a key design criterion for many applications. In the present study, the focus will be on drive cycle analysis and the required operating points are all well within the limiting torque curve of the engine. Under these conditions, the problem is reversed: to achieve a specified power output under more humid conditions, additional fuel is required. In extreme conditions this is seen when analysing studies of intake gas water

injection used in prototype setups for NO_x emissions control [137]. In the case of a drive cycle it is difficult to see how the correction factor may be applied as this would need to adjust the operating point of the engine during the test. To the author's knowledge there are no current correction factors for fuel consumption due to this effect, probably because the effect is small and not problematic in previous studies.

It is known that ambient conditions act as a disturbance to the engine behaviour. Some correction factors will be applied for NO_x emissions however there are none existent for fuel consumption. Full conditioning of combustion air would be a solution but involves large investment in facilities. It is recommended that this suspected behaviour be considered when comparing experimental results and that in future should this become a limiting factor, experimental campaigns be carried out on controllable rigs to establish such a correlation for future use.

3.2.2. Process control in the test cell

Detailed work relating to testing repeatability has been carried out on a chassis dynamometer facility [96, 97]. The work was related to aspects such as vehicle alignment, tyre pressures, dynamometer speed levels, vehicle cooling and battery state of charge. This work used measured data from tests with purposely perturbed conditions to establish setup tolerances for vehicle installation to meet a required level of repeatability. A number of the factors investigated in this work would not be relevant to engine dynamometer testing because they relate to the vehicle rather than the engine, however it is worth discussing aspects relating to the battery and cooling.

Battery voltage is known to be a poor measure of state of charge experiments showed poor correlation between battery voltage and measured fuel consumption [96]. Assessing energy consumption over the drive cycle gives a more meaningful measure: Battery current and voltage can be measured over the drive cycle and used to give a measure of energy use over the test. This was performed for different initial states of charge good correlation with fuel usage. While this method was acceptable for identifying poorly charged batteries, it is only applicable after a test has been conducted, meaning the test will be wasted before the operator is informed. This highlights the need for a rigorous battery management to ensure a good state of charge before each test.

The thermal state of the engine was also assessed; however surprisingly initial results looking at the effect of start temperature on fuel consumption showed a very small effect. Further investigations showed that a stronger correlation was obtained between fuel consumption and cumulative or mean temperature over the experiment. A number of experiments with varying cooling fan speed and bonnet opening showed that whilst there was good agreement between temperature and fuel consumption, the control of engine cooling was fairly poor [96].

4. *Assessment of facility capabilities*

4.1. Statistical power test

The concept of demonstrating statistically significant benefits in fuel consumption was presented at the beginning of this chapter. Before embarking on an experimental campaign it is possible to estimate how many tests will be required to demonstrate the expected difference. This method is called the statistical power testing [138]. The power of a statistical test is given as a number between 0 and 1 with a higher power meaning there is less chance of incurring a type II error, i.e. not finding a significant difference when there truly is a difference. In medical statistics a test should have a minimum power of about 0.8. For engine testing, to the author's knowledge, no expert review has been published on this topic. The power of a test is dependent on the following factors:

- **Required significance level**: reducing the significance level will increase the power, but this also increases the likelihood of a type I error, i.e. finding a statistically significant difference when in fact the batches are the same. Significance levels are usually prescribed by customer demands.
- **Effect size**: "Large effects are easier to find". More specifically, it is the size of the effect relative to the population variance that is important.
- **Measurement precision**: Improving the measurement precision will reduce the variance of that data. Steps described in the previous section will contribute to this.
- **Design of experiments**: Number of tests and optimally setting the independent variables can increase the statistical power (optimal test design increase the power for calculating model coefficients).

Performing power analysis *a priori* (before data collection) is useful to determine the number of tests, the required measurement precision or the smallest detectable effect. The power is defined as 1-probability of a type II error and the concept is illustrated in figure 3.20. The distributions of results from two measurement samples are shown, for example and baseline and a perturbed setup. The power of the significance test is the proportion of tests from sample two that would be judged significantly different from sample 1 using the statistical test. Further details relating to the theory are beyond the scope of this thesis and for further information the reader is directed to [138].

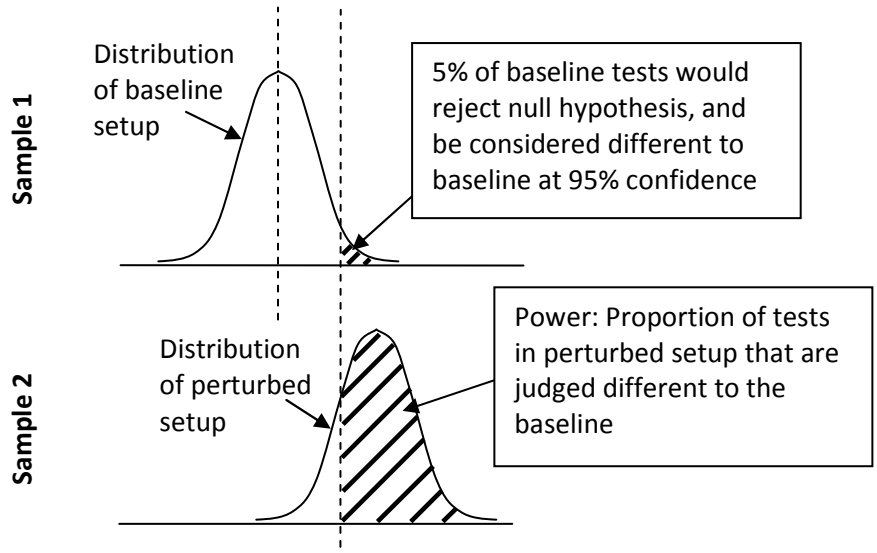


Figure 3.20: Graphical illustration of the power of statistical tests based on measurements from two distinct experimental setups

4.2. Application to experimental facility at Bath

A series of 14 repeat cold-start NEDC tests were performed on the facility to determine the performance of the measurement systems. For the given setup, the mean and standard deviations for the 14 test samples are given in table 3.1.

	NO_x	Gravimetric FC	Carbon Balance FC
Mean (g)	4.95	892.84	933.04
Standard deviation (g)	0.09	3.67	4.14
Coefficient of variation	1.8%	0.4%	0.4%

Table 3.1: Statistics for selected measurements from 14 repeat cold-start NEDC tests

Statistical power analysis has then been applied to determine a number of performance criteria for the experimental facility and are summarised in table 3.2. These results clearly show that the fuel consumption measurement processes perform better than the NO_x emissions. For example series of 6 repeat tests would demonstrate a 0.6% change in fuel consumption but only a 2.7% change in NO_x.

	NO_x	**Gravimetric FC**	**Carbon Balance FC**
Number of tests to determine 2% change	9	3	3
Number of tests to determine 1% change	30	4	4
Number of tests to determine 0.5% change	111	8	9
Min observable change for 3 tests	6.1%	1.3%	1.5%
Min observable change for 6 tests	2.7%	0.6%	0.6%
Min observable change for 12 tests	1.7%	0.4%	0.4%

Table 3.2: Statistical power test performance criteria for the experimental facility at Bath

These results can be used for experimental planning when the some estimate of the difference that needs to be demonstrated is known. In this thesis, this estimate will be a result of response models based on empirical data from designed experiments but in other cases this could be obtained from purely analytical simulations. Clearly this estimate of required experimental effort will be useful for busy experimental facilities and budgeting exercises.

This approach is less useful in determining the number of experiments in multi-dimensional analysis, such as the DoE approach. In this case, the analysis suggested that the complete DoE matrix should be repeated a number of times. In most cases, this will be impractical because of the time implications and facility availability. Bayesian statistical approaches aim to include prior knowledge of the system behaviour and system variance to reduce the number of data points required for a model of given quality. It would be interesting to investigate an approach somewhere between this Bayesian approach and the statistical power analysis; in this case the actual system response would not be known, but an estimate of the measurement capability would. This could then contribute to the experimental design phase in determining a suitable test matrix to explore the system response to various inputs and achieve a response model of required quality.

5. *Chapter summary and conclusions*

In this chapter improvements to the experimental facility for demonstrating small differences have been presented, tested and quantified. The design of experiments approach was reviewed and is established as an industry standard. Improvements to the accuracy and repeatability of fuel consumption and emissions measurements were presented. Finally the measurement capability or the facilities was presented based on statistical power analysis. Based on the findings from this chapter the following conclusions may be drawn:

1) The DoE approach will be essential to reduce the number of test points when performing cold-start experiments because of the lengthy thermal soak periods.

2) State of the art DoE methods for capturing the dynamic engine behaviour remain in infancy and have so far not captured engine behaviour during warm-up.

3) Although highly accurate measurement devices are commercially available for measuring fuel consumption, the integration of these devices into the experimental facility is a source of inaccuracies. The inclusion of physically based correction factors can improve this accuracy and provide statistical agreement between independent methods.

4) Ambient conditions appear to impact NO_x to a greater extent than is expected in well-known correction procedures. A linear trend appeared between NO_x emissions and humidity, meaning an empirically derived correction factor was derived to allow tests with significantly different ambient conditions to be compared. However, the limited data means this correction factor should be used with caution.

5) Statistical power analysis suggested that to demonstrate a 0.5% change in fuel consumption over the NEDC, eight repeat tests would be required. The system appeared less capable for NO_x emissions with eight repeat tests capable of demonstrating a 2.5% change.

Chapter 4 - *Experimental facilities, Prototype Hardware and testing procedures*

Having reviewed the state of the art in terms of parasitic loss reduction in chapter 2, this chapter will describe the candidate hardware and engine configurations to be used in this study. The experimental approach and necessary instrumentation for demonstrating these changes will also be presented. Engine hardware was chosen in collaboration with project partners with the aim of delivering performance benefits with low cost impacts. By adopting this pragmatic approach there is a larger potential for commercial adoption.

In a first section, an outline of the experimental approach will be presented. Section 2 will then give an overview of the experimental facility and engine architectures. This has been discussed in part in chapter 3 for fuel consumption measurements, however a more general perspective will be taken. Section 3 will describe the candidate hardware and its integration into the engine system. In section 4, details relating to the experimental campaign will be presented. Finally specialised instrumentation and data processing will be presented in section 5.

1. *Approach and methodology overview*

In a first stage designed to provide baseline data, the fundamental physical processes will assessed using a production engine, stripped of all of its auxiliary systems and set up to give direct control of the engine inputs. This was used for steady-state testing to assess the influence of engine temperature on various aspects of engine behaviour. This was performed in an environment where full control of engine cooling fluid temperatures was available from the experimental facility.

Secondly, the dynamic behaviour of the production engine system will be assessed over the NEDC[35] with a number of concept hardware devices designed to alter the engine behaviour during warm-up. This will allow the performance of each hardware device to be assessed as well as their impact on engine output parameters such as emissions and fuel consumption

With an understanding of the fundamental behaviour of the engine and new system hardware, model based approaches will be used to understand the interactions between conventional control parameters and the additional control parameters introduced with the new hardware. With a detailed understanding of the new system, optimised calibrations will be determined and validated.

Finally a dynamic DoE approach will be used on the stripped engine (without additional cooling system controls) for fundamental modelling of engine emissions from cold-start. This aims to develop a global empirical model for engine performance. Because the engine temperature cannot be modulated rapidly as is the case with conventional calibration variables[36], a simple approach will be presented for capturing this temperature dependent aspect.

2. *Engine hardware and experimental facilities*

Two similar high-speed, direct-injection, turbocharged and intercooled Diesel engines were used for all experimental work. The main differences were swept volume and the fuel injection equipment manufacturer. The first of these two engine was a Ford Puma 2.0L and was

[35] This duty cycle will be presented later in this chapter.

[36] Although the stripped out engine facility enabled controlled variation of engine cooling water, there are distinct differences in the temperature distributions of an engine during warm-up and a warm engine with cold coolant flowing through the water jacket.

installed on a transient dynamometer with minimal production accessories. This allowed easier control of the engine behaviour for understanding the core physical processes. The second engine was a Ford Puma 2.4L engine of similar age, design and control strategy. This engine was also installed on a transient dynamometer, but was equipped to near production specification to assess the inclusion of new hardware. The next sections provide an overview of the two engines and test facilities.

2.1. Ford PUMA 2.0L engine

The 2.0L PUMA engine was a EURO IV specification[37], turbocharged Diesel engine. It was equipped with a common rail fuel injection system capable of delivering 1600bar rail pressure. The engine was fitted with a coolant cooled, high pressure EGR loop. The engine was installed on a David McClure 215kW transient dynamometer facility with the gearbox removed and replaced with a direct drive. A picture of the installation is shown in figure 4.1 and a diagram of the coolant circuit layout is shown in figure 4.2.

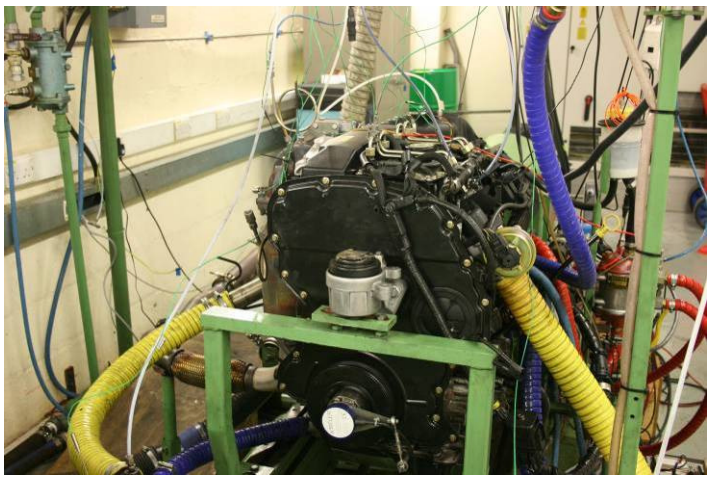


Figure 4.1: Photograph of 2.0L PUMA experimental facility

On this facility the majority of engine auxiliary devices were removed. The external cooling circuit was replaced by a coolant/water heat exchanger to directly control coolant temperature. Similarly, the intercooler was replaced by an air/water heat exchanger and again could control post inter-cooler gas temperature. EGR cooling was still performed by the main engine coolant. It is important to note that such a system allows the engine to operate well outside what is feasible in-vehicle where cooling is provided by air flow.

[37] EURO IV refers to the legislation on harmful emissions within the European Union. For light duty commercial vehicles, EURO IV refers to requirements for vehicles certified between 2005 and 2009 [5].

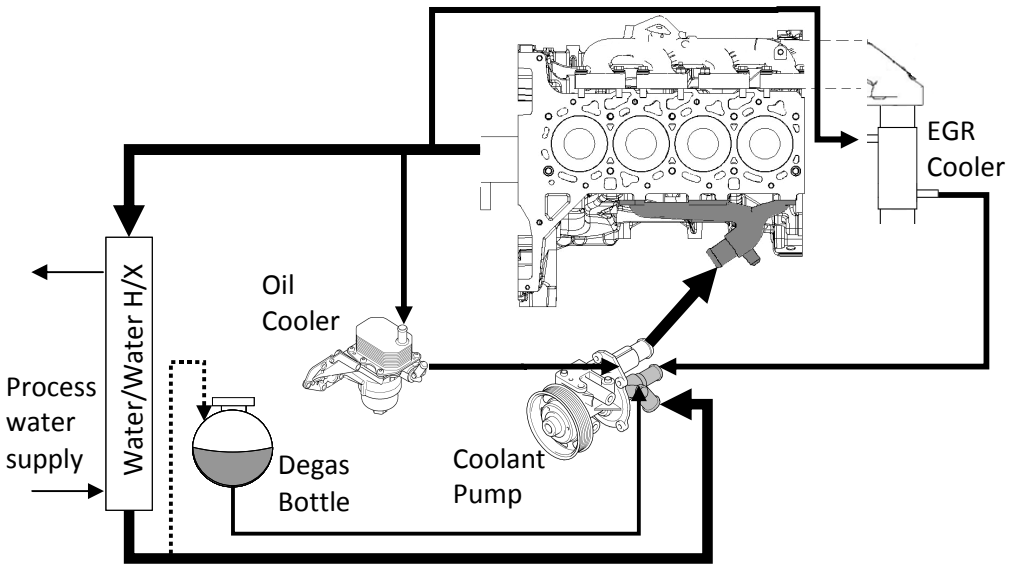


Figure 4.2: Diagram of engine cooling circuit for 2.0L PUMA engine

2.2. Ford PUMA 2.4L engine

Two series of tests were performed with this engine. The first was aimed at proving the viability of new hardware on the engine. The second phase involved transferring this hardware to an identical, but heavily instrumented version of this engine for in depth analysis. In this section the general features of the engine will be presented and the additional instrumentation used for second phase will be described in section 5. As this work focussed on the lubricant and cooling circuits, the baseline circuits will be described as modifications and analysis will require a basic understanding of these layouts.

The 2.4L PUMA was also designed to a EURO IV specification and installed on a similar dynamic facility. As with the 2.0L engine, the gearbox was replaced with a direct drive to the transient dynamometer. A photograph of the experimental facility is shown in figure 4.3. The engine had a similar specification including turbocharger, intercooler and common rail fuel injection, and in this setup the majority of the production auxiliaries were also installed.

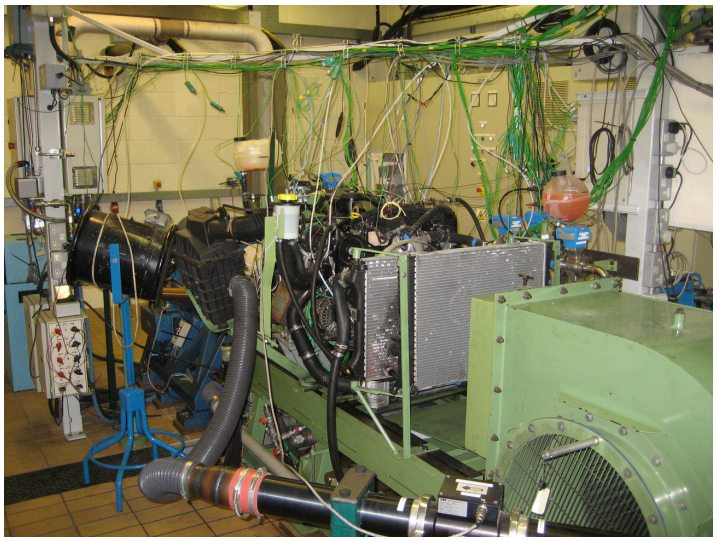


Figure 4.3: Photograph of PUMA 2.4L experimental facility

The production engine cooling system was replicated with the exception of the cabin heater matrix. This was removed because the focus of this work was to improve fuel consumption and emissions performance. Cabin heating is not used during this aspect of the homologation procedure[38]. A diagram of the production cooling circuit is shown in figure 4.4 and is typical for engines of this age, with the exception of the Pressure regulated thermostat (PRT)[39]. The coolant circuit is also required to cool EGR gases, the EGR valve and engine lubricant. A large electric fan was installed to provide cooling air flow over the engine radiator and simulate vehicle movement. The fan speed was calibrated to give representative top-hose coolant temperatures.

Figure 4.5 shows a simplified layout of the production lubricant circuit. Oil is pumped from the sump through the oil cooler and filter assembly and up into the main oil gallery. The main bearings and piston cooling jets are fed directly from this gallery. From the main bearings, oil is fed through the crankshaft to the big-end bearings. The gudgeon pins, piston rings and piston skirt are all splash lubricated. Another feed from the gallery allows oil to flow to the head assembly to serve the valve train. The flow splits between intake and exhaust sides and flows through the first bearing in each cam shaft before being distributed to the other 4 bearings. All

[38] Improvements to devices that are not tested in the homologations cycles but that positively affect real-world fuel consumption can receive “Eco-innovations credits” of up to 7g/km in the European Union [5].

[39] Like a traditional thermostat, this component is still based on a wax element, however it is sensitive to both top hose and bottom hose coolant temperatures meaning that when a large cooling potential is available over the vehicle radiator the thermostat will have a tendency to stay closed even if top hose temperatures are hot [23]. This device was demonstrated to offer superior coolant temperature control to the standard production thermostat during a preceding product and was retained for this study.

oil leaking from bearings in the head is collected and flows down return passages to the sump; oil leaking into the crankcase from the main bearings and sprayed by the cooling jets returns directly to the sump. There is also a direct oil return from the turbocharger to the sump.

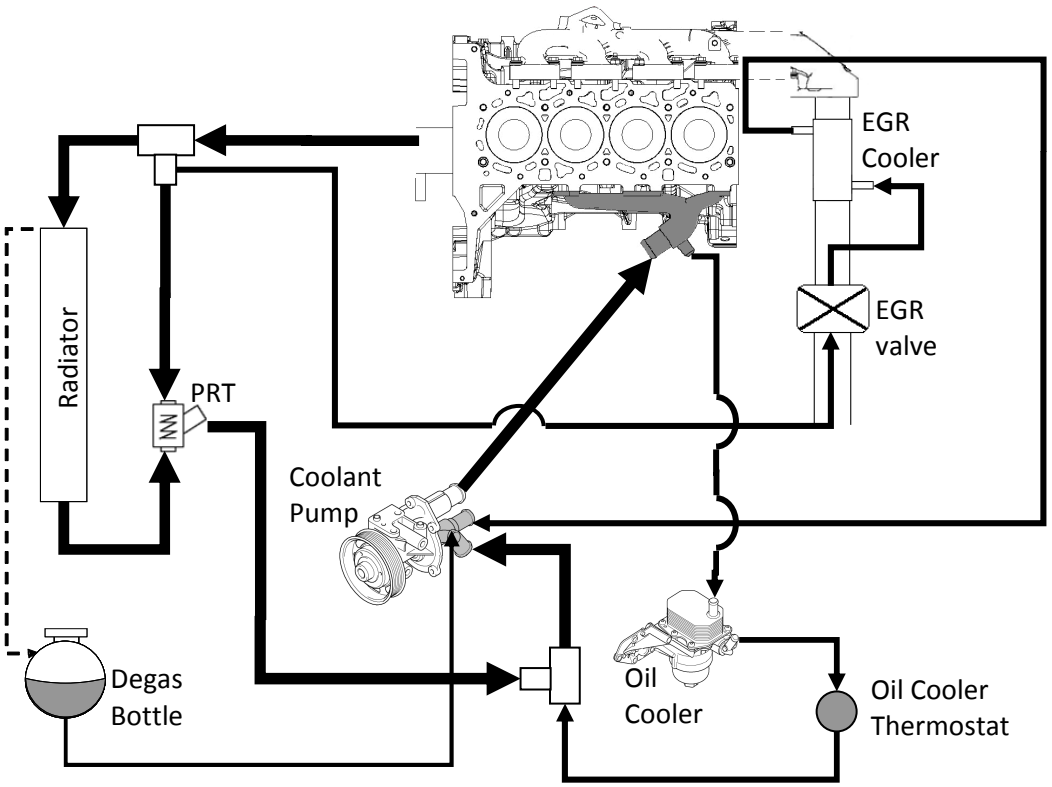


Figure 4.4: PUMA 2.4L production cooling circuit with heater matrix removed

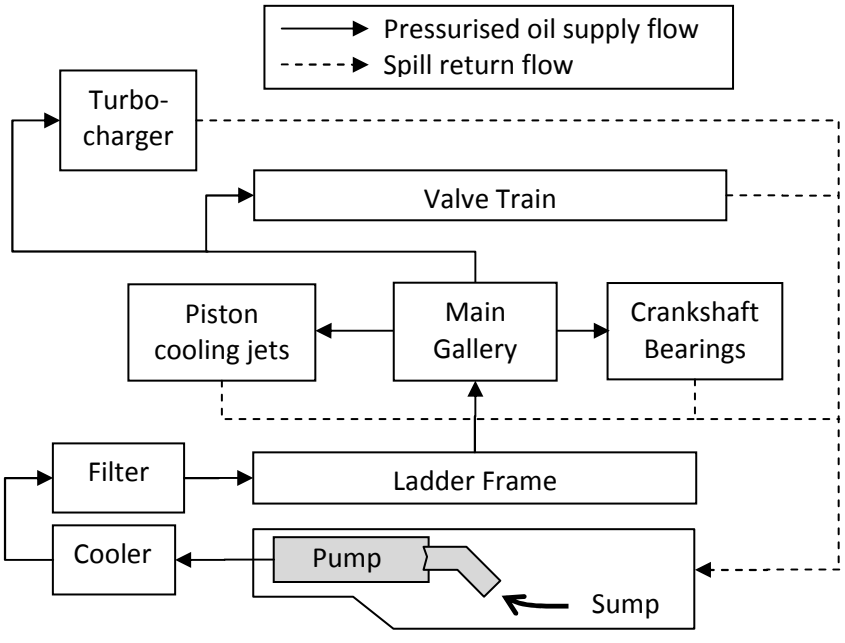


Figure 4.5: Simplified diagram of PUMA 2.4L production lubricant circuit

3. *Candidate hardware*

3.1. Hardware description

In an attempt to improve the warm-up rate of the engine a range of candidate hardware was sourced and integrated into the engine. These devices were coolant flow control valves, a dual EGR system and a variable flow oil pump. The following sections provide details for each of these. In the subsequent section, the integration of these devices into the production circuit will be discussed.

3.1.1. Coolant flow control valves

Simple butterfly valves were used to control the flow of coolant in different legs of the circuit. The valves were actuated by production rotary electric actuators (REAs) typically used to modulate VGT position. The REAs were controlled by the engine ECU. This integration with the ECU allowed the valves to be mapped against other ECU inputs which increased the degrees of freedom of calibration. It was decided that the valves should be controlled using a mapped approach based on engine speed, fuelling quantity and cylinder head metal temperature, measured using a production sensor. This approach, similar to some reviewed in chapter 2 [19, 33], was preferred over coolant temperature control schemes. When the coolant is within its working range, its temperature is not the primary area of interest, rather it is the thermal state of the engine itself which is the object of the control scheme.

Directional flow control valves were also used integrated into the control system in a similar way to allow components to be removed from or included in the circuit easily via the control system.

3.1.2. Dual EGR cooler

A dual EGR cooler system was installed which allowed modulation of heat flows within the engine. One of the coolers used coolant as in the production setup, whereas the second cooler used oil (see figure 4.6). The second EGR cooler was attached to an extended exhaust manifold and was a duplicate of the first, using the same production hardware. A photograph of the arrangement is shown in figure 4.7. Note that the EGR valve is of the liquid cooled type, and both valves were connected in parallel to the coolant circuit. On the liquid side, there was no

provision to stop either oil or coolant flow when that particular EGR leg was not in use, meaning that both oil and coolant were flowing in their respective coolers at all times.

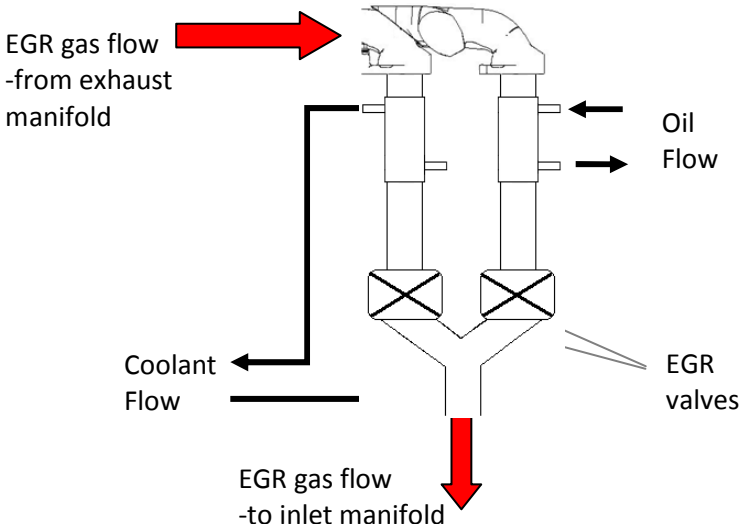


Figure 4.6: Dual EGR cooler setup

Although on the gas side the EGR coolers were arranged in parallel, the control of the EGR valves was not designed to use both coolers simultaneously. The EGR cooler changeover was accomplished by maintaining one of the EGR valves in the closed position while the other followed the duty cycle determined by the production engine control algorithm. Operating both coolers simultaneously was not planned since the dual setup was simply an expedient approach to allowing a flexible evaluation of the merits of each cooling medium, not a production intent system. Any production implementation would use one or the other system, informed by the findings of the research. It could be argued that increasing the level of EGR cooling by using both coolers simultaneously offers additional benefit by altering the NO_x response of the engine, but this was not the primary objective of the research undertaken in this thesis.

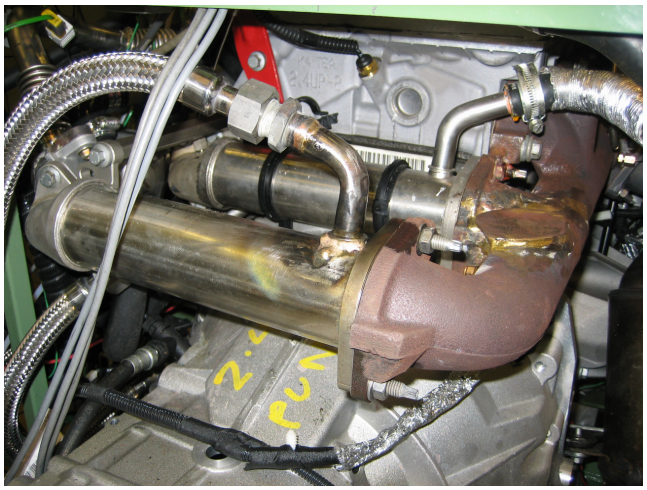


Figure 4.7: Photograph of dual EGR cooler hardware

3.1.3. Variable Flow Oil Pump (VFOP)

A variable flow oil pump was sourced to replace the production device. The production oil pump is a fixed displacement, off-axis design and the variable flow device is sufficiently similar to replace this with minimal hardware changes[40]. Photographs of both pumps are shown in figure 4.8. The pump displacement was controlled electro-hydraulically using an algorithm resident within the ECU which sought to adjust oil pressure as a function of engine speed and load (as represented by fuelling quantity). The ECU control acted open-loop, with hydraulic feedback to the pump controlling displacement volume to provide a stable gallery pressure.

The pressure set-point maps used within the ECU were adjustable via the Ati Vision *no hooks* calibration tool. In this way a stand-alone passive hydro-mechanical controller was combined with active electronic control to produce a robust yet adaptable system. Clearly a fail-safe design is desirable in a system so critical to engine health as the oil pump. Mechanically simpler electronic control schemes using direct control of oil pressure do not offer this capability and their operation in the case of a motoring engine and/or electronics failure must be carefully considered. Prior to this project, a "lowest permissible pressure" calibration was set up using knowledge derived from expert engine developers. Other simple calibrations were developed during this work to provide a constant pressure over the speed load range[41] or maximum pump displacement. The aim of these calibrations was to investigate the relationship between oil pressure and fuel consumption.

[40] A pressure tap drilled to the ladder frame for oil pressure feedback to the pump.
[41] This value was constant wherever engine speed allowed as at low engine speeds the maximum pumping pressure can be lower than the desired set-point.

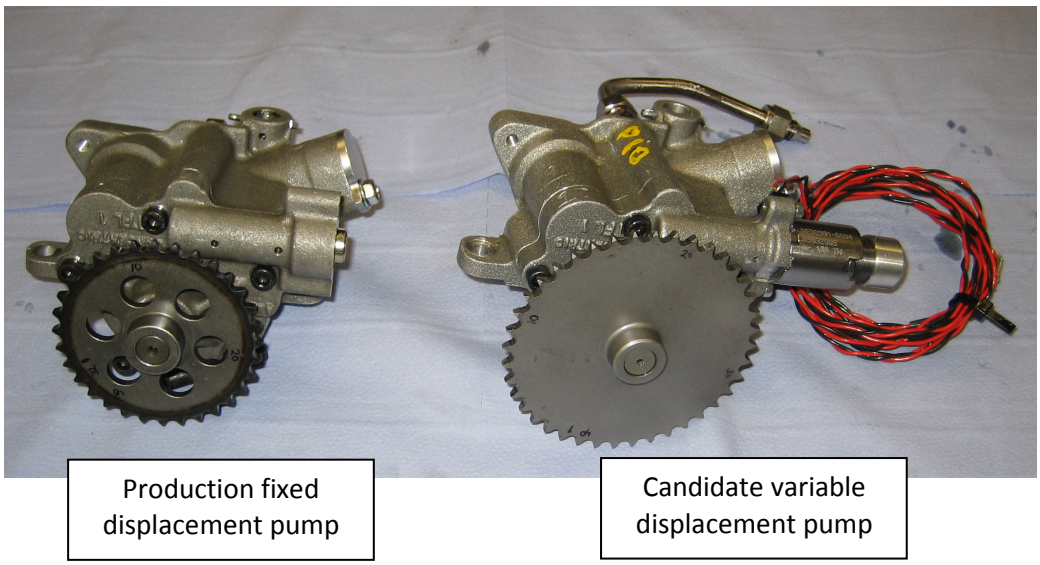


Figure 4.8: Photographs of fixed displacement production oil pump and variable flow oil pump

3.2. Hardware Integration

Each of the prototype circuits detailed in this section were tested on the PUMA 2.4L engine over the NEDC. Basic changes to the engine behaviour during warm-up can be achieved by changes to the circuit without any hardware addition, however to move towards an active thermal management system the integration of the above candidate hardware is necessary.

3.2.1. Basic changes to engine TMS

The production engine incorporates a thermostat to control coolant flow through the oil cooler. The design is intended to avoid oil cooling during warm-up, but then allow heat to be evacuated from the oil under hotter conditions. In practice both at the facility at the University of Bath and results from the literature, coolant temperatures lead oil temperatures during warm-up and it would be beneficial to allow heat transfer between coolant and oil, in effect turning the oil cooler into an oil heater[42]. The removal of the thermostat in the oil cooler leg of the coolant circuit would allow this and transform the functionality of the oil cooler into a more general oil/coolant heat exchanger.

[42] Clearly for a production vehicle the impact on cabin heating would also need to be assessed.

3.2.2. Active TMS

The prototype active thermal management system designed and installed on the PUMA 2.4L engine made use of the following hardware:

- The Dual EGR cooler system
- Two coolant throttles installed in different legs of the cooling system
- Flow control valve to include or isolate the oil/coolant heat exchanger
- Variable flow oil pump

The changes to the external oil circuit and to the coolant circuit are shown in figures 4.9 and 4.10 respectively. The external oil circuit now incorporated an EGR cooler in addition to the oil filter cooler assembly. A small adaptor plate was also required to divert the flow to and from the production parts. The external oil circuit increased the required oil volume by approximately 0.6L (10%) and the estimated mass of the additional components was 3kg.

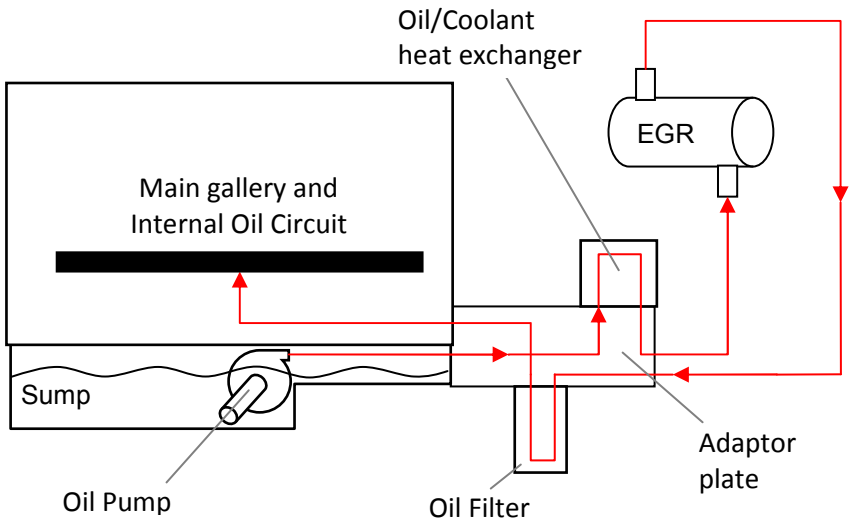


Figure 4.9: Modified external oil circuit showing oil flow

The modified coolant circuit included two coolant throttles, a flow control valve and the additional EGR valve associated with the dual EGR cooler system. The plumbing was also modified relative to the production system to encourage heat transfer to the engine oil during warm-up. The oil cooler was no longer plumbed into the colder coolant on the inlet manifold, but downstream of the EGR cooler where coolant is expected to be hottest following heat transfer in the engine and from the EGR gases. The first coolant throttle (a) was installed to

restrict flow in the main circuit at the front of the engine. Like Choukroun and Chanfreau [34] and Torregrosa *et al.* [43], this throttle may provide isolation of the front engine coolant circuit during warm-up. The second throttle (b) was installed to restrict flow in the EGR loop. The flow control valve (d) was installed to include or isolate the oil cooler (c) from the main coolant circuit, effectively replacing the passive thermostat in the production system by an active device. The increase in coolant volume due to the additional hardware was minimal as they were installed in place of hoses from the previous build[43].

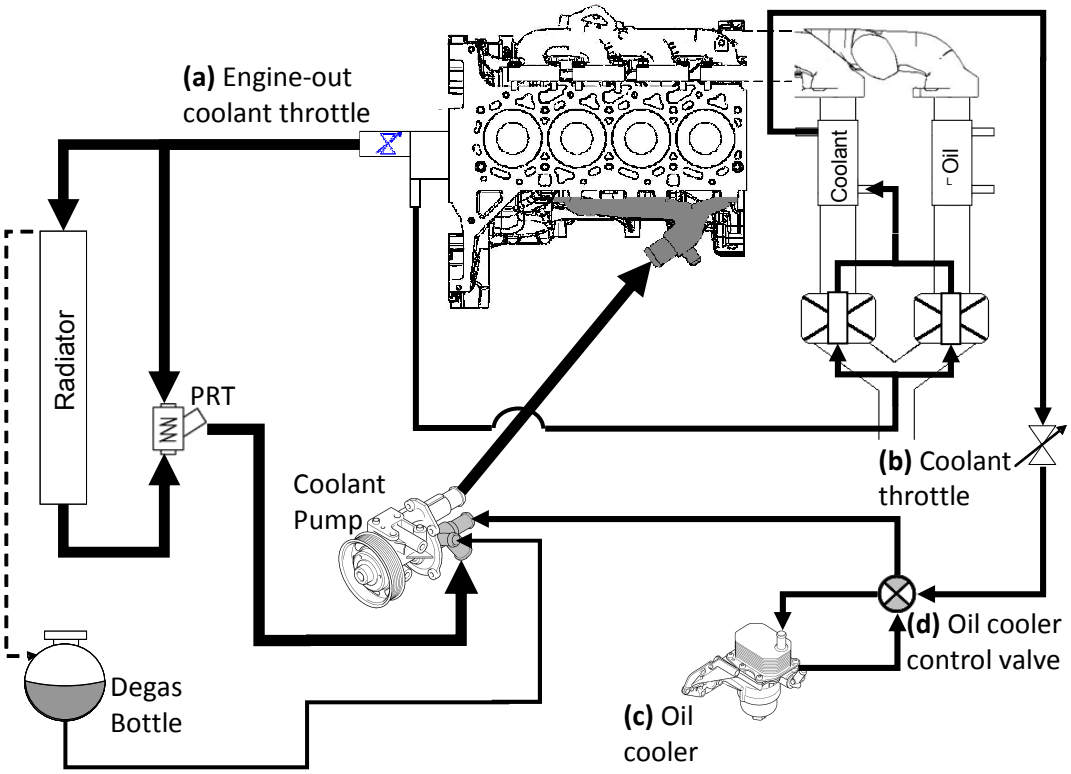


Figure 4.10: Modified coolant circuit

[43] The additional hardware did not cause any significant increase in the coolant volume for the circuit with flow meters installed. However, the integration of flow meters into the coolant circuit caused a significant increase in coolant volume (2L or 25%) over the production system. It is certain that additional hoses for the flow meters eased the fitting of flow control valves without significant increase in volume and the fitting of these devices into a production circuit would inevitably cause an increase in coolant volume.

The design for this circuit and control actuators was aimed at allowing modulation of heat flows during engine warm-up. The modifications to the oil and coolant circuits allowed control of the following parameters:

- Control of engine-out coolant flow via throttle (a)
- Control of coolant flow in EGR cooler circuit via throttle (b)
- Control of coolant flow to oil cooler via control valve (d)
- Control of oil pressure using VFOP
- Use of coolant- or oil-cooled EGR via use of appropriate EGR valve
- Other common calibration parameters

4. *Experimental programme*

The project was organised into the following three steps that logically build on each other, progressively applying a more complete representation of the engine behaviour. These three steps are steady state tests, dynamic NEDC tests and global dynamic modelling. Each of these will be presented in the following sections.

4.1. Steady state testing

A number of steady state tests were performed on the PUMA 2.0L engine where coolant temperature was artificially varied using the test control system (see figure 4.2). These tests aimed to understand the basic engine behaviour under different thermal conditions. This testing approach is useful in that the complex behaviour of the engine under transient conditions is avoided. On the other hand, the operating conditions will be substantially different to those during warm-up. The artificial modulation of coolant temperatures will result in different thermal distribution compared to warm-up conditions.

The steady state experiments were performed at key operating points on the NEDC and two points were chosen representing a 50km/h and 120km/h cruise. These were chosen to represent a low and high load point, as differences were seen in the literature between these two conditions [23, 89]. At each speed/load point, the interactions between engine thermal state, injection timing and EGR rate were established using a full factorial DoE approach.

4.2. NEDC tests

Cold- and hot- start NEDC tests allow the dynamic behaviour of the engine to be captured along with realistic warm-up thermal conditions. This cycle offers a good platform to assess the performance of the candidate hardware to improve warm-up and is directly transferable to vehicle homologation. As prescribed by the legislative cycle, experiments from cold start were conducted from a start temperature of 25°C[44]. However it is important to note that his is not representative of real world conditions where cold start temperatures would be much lower. At these colder temperatures the prospective benefits in fuel consumption from improved warm-up would be expected to be larger as the oil viscosity further increases. The PUMA 2.4L engine was taken from a light commercial vehicle; however on the engine test stand the powertrain between the engine and wheels is simulated by the dynamometer. The duty cycle for the PUMA 2.4L engine is shown in figure 4.11: as the NEDC is specified by vehicle speed, the engine speed and torque traces were calculated using CP engineering VSIM software (Vehicle simulation). These are subsequently specific to this engine, gearbox and vehicle configuration.

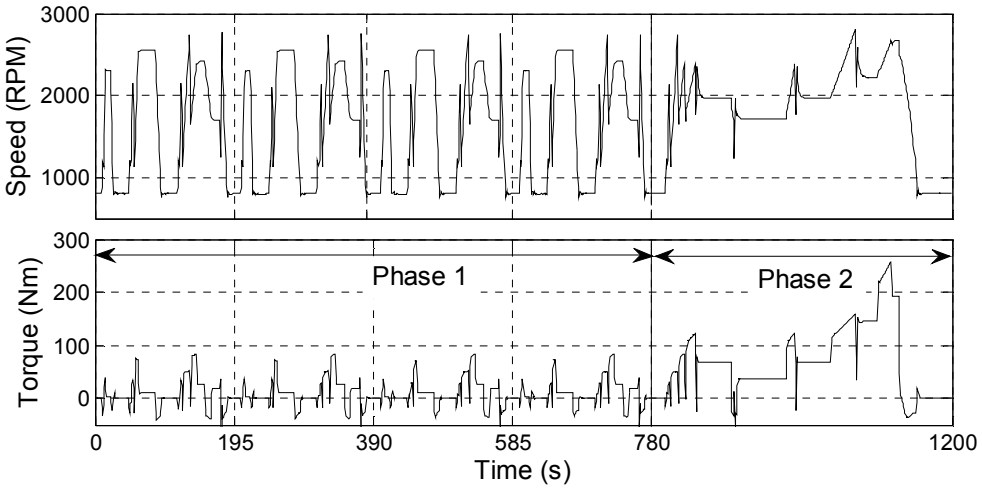


Figure 4.11: Speed and Torque for NEDC cycle for 2.4L PUMA engine

Figure 4.11 also shows a common split into two phases after 780 seconds between the low power urban cycles (ECE15) and the higher power extra urban cycles (EUDC). This split is convenient as on the installation at the University of Bath, phase 1 approximately represents the coolant warm-up period, with the thermostat opening around the start of phase 2. Each of the NEDC tests in this work will be run from either cold- or hot-start conditions. In both cases a

44 A -7°C cold start test is in force for gasoline engines aimed at CO and HC emissions legislation.

rigorous temperature soak phase is performed before the cycle. For cold-start tests this is either an overnight soak at 25°C or a proven forced cool down procedure [100]. For hot-start tests this is a 20 minute conditioning phase[45].

Initially a scoping exercise was performed aimed at understanding the behaviour of both the production and active thermal management systems. The results from this initial investigation were combined with those observed from the steady state testing to establish the pertinent inputs and input ranges to acquire in depth understanding of the system and interactions. NEDC tests were subsequently performed in a structured way: a DoE approach was used whereby each test point was in fact an entire drive cycle, with appropriate adjustments to the control strategy as dictated by the experimental design. Each of these cycles was run over the NEDC cycle from a cold-start condition. In this way an experimental design approach more commonly applied to steady state test points was applied to a dynamic test cycle, with the experimental factors applied throughout each cycle to allow the effect on both final results and dynamic behaviour to be studied. The resulting knowledge was used for system optimisation applicable to the drive cycle.

4.3. Dynamic Engine modelling

The NEDC tests presented above, in combination with the DoE approach are expected to give a good appreciation of engine behaviour. However, the knowledge gained from those experiments will be limited to that particular duty cycle. With different legislation for different regions of the globe, it is a clear benefit if a global, universally applicable approach can be used. The concept of dynamic design of experiments was introduced in chapter 3 and this methodology was applied to the PUMA 2.0L engine. Although this compromises studying the effect of the active TMS, it did allow the method to be demonstrates within the time frame of this thesis. The dynamic modelling approach is a powerful tool, but the effort required to drive the engine through the required training sequences should not be underestimated.

The synchronised control of each of the desired set-points is critical both to ensure that the training data covers the desired design space and to ensure the engine remains within safe and sensible operating limits. The communications layout is summarised in figure 4.12: the experiments were controlled by the host system running CP Engineering Cadet software: this

[45] To ensure repeatable thermal soaks for hot-start NEDC cycles the engine was run for 20min at a simulated 70km/h cruise after thermostat opening before the drive cycle.

directly controlled the dynamometer speed and accelerator pedal position. An ASAP 3 link[46] was installed between the host system and a second computer running the ATI Vision calibration tool (CAN tool). The calibration tool in turn communicated with the ECU using the industry standard CAN calibration protocol (CCP) [139]. ECU variables were controlled by the host system using this link. Other than ECU variables, all channels were recorded directly to the host system. The ECU variables were recorded within the CAN tool as the ASAP 3 connection did not offer sufficiently high update frequency. Engine speed was recorded on both systems and subsequent post processing was required to align the two data sources.

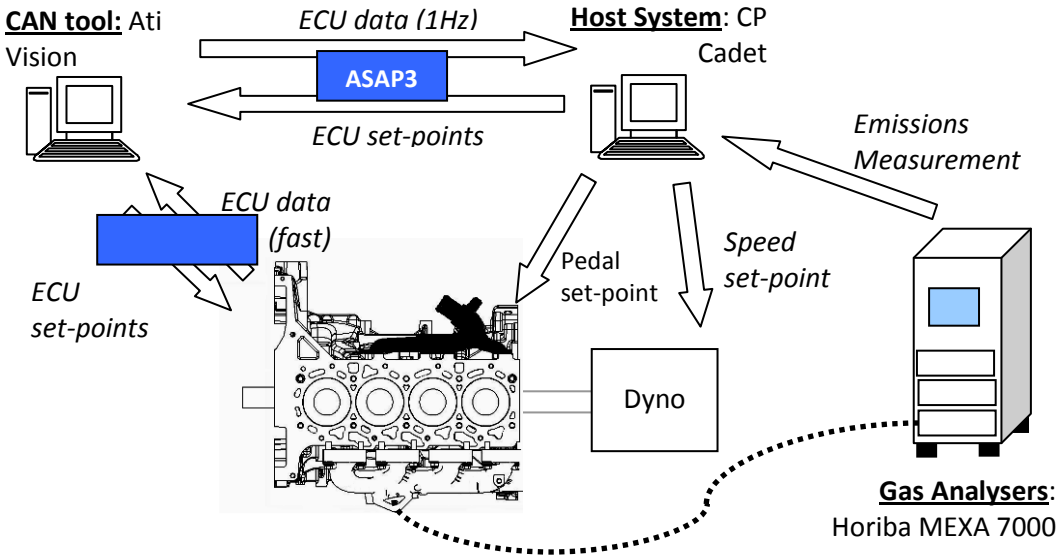


Figure 4.12: PUMA 2.0L test cell communications

5. *Instrumentation and Measurement systems*

5.1. Emissions analysis

The emissions analysis and carbon balance fuel consumption have been described in chapter 3 where the accuracy of these processes was assessed and improved. Emissions concentrations were measured continuously for both pre- and post-catalyst gas samples wherever appropriate using two Horiba MEXA 7000 analysers. In addition the CO_2 concentration in the inlet manifold was measured to calculate EGR rate. The measurement for each emissions species was carefully time aligned [125, 126] for each test to account for the transport delays

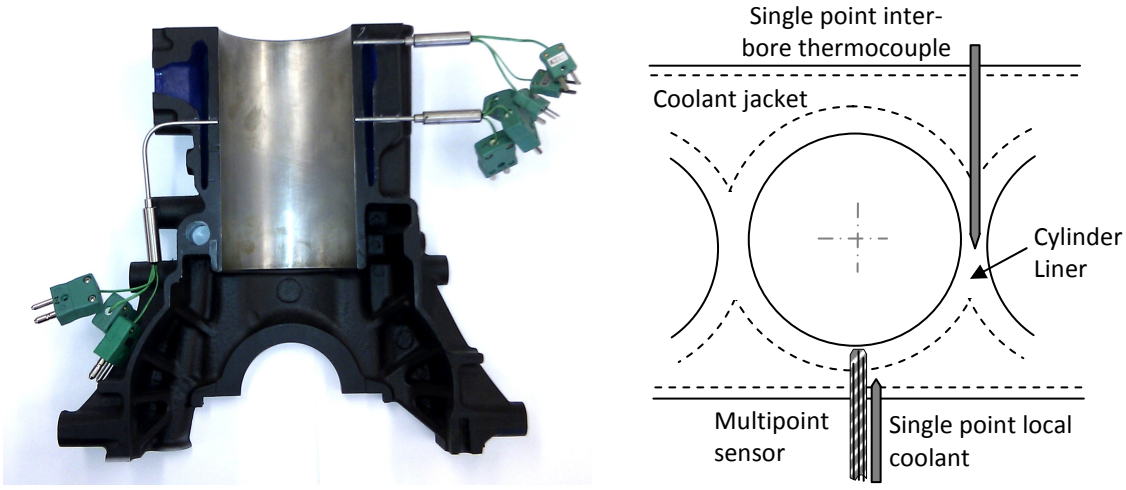
[46] An ASAP 3 link is an industry standard communications system for data exchange between a test cell host system and calibration tool [139].

between the exhaust line sample port in the test cell and the analysers in a common analyser facility. In line with the findings from chapter 3, all emissions measurements by mass were estimated according to the appropriate British Standard [127].

5.2. Heat flux analysis

5.2.1. Instrumentation

The second iteration of the 2.4L PUMA engine was fitted with over 60 thermocouples within the engine structure. A large proportion of these were situated around the combustion chambers of cylinders 2 and 3 while the others were located in the internal oil circuit to measure local oil temperatures. Figure 4.13 shows both a photograph and diagram of how the three point sensors were installed in the engine block to measure the temperature distributions through the cylinder liner. In each case the depth from the combustion chamber wall was carefully controlled during machining. Where the multipoint sensors were fitted in proximity of the coolant jacket, an additional thermocouple was installed to measure the local water temperature to estimate convective heat transfer coefficients. Figure 4.14 shows the location of thermocouples around the cylinder liners. Some of the thermocouples are single point measurements such as in between the cylinder bores while others are grouped into triple-point sensors to measure thermal gradients in the liner.



(a) *(b)*

Figure 4.13: (a) model of three point thermocouple installation in cylinder liner and (b) diagram of installation of thermocouples

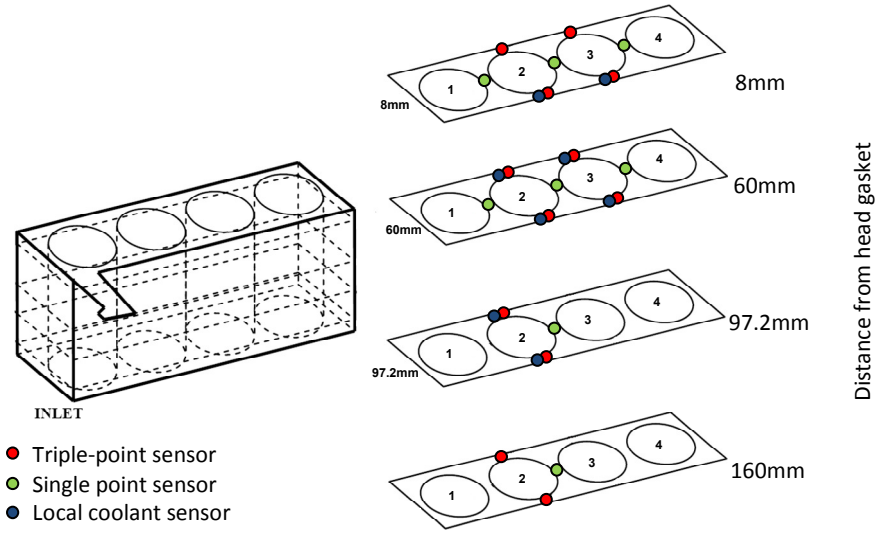


Figure 4.14: Engine block thermocouple locations

Figure 4.15 shows the layout of the multipoint sensors: these consisted of three 1.5mm thermocouples with the junctions set 2mm apart, packaged within a 3mm diameter stainless steel sheath, backfilled with magnesium oxide to ensure good thermal contact. The position of the thermocouples relative to each other is of key importance for an accurate estimate of the temperature gradient. An extensive calibration procedure was undertaken on custom made calibration equipment to ensure these gave accurate results. The details of this calibration procedure have been published by Lewis *et al*. [140].

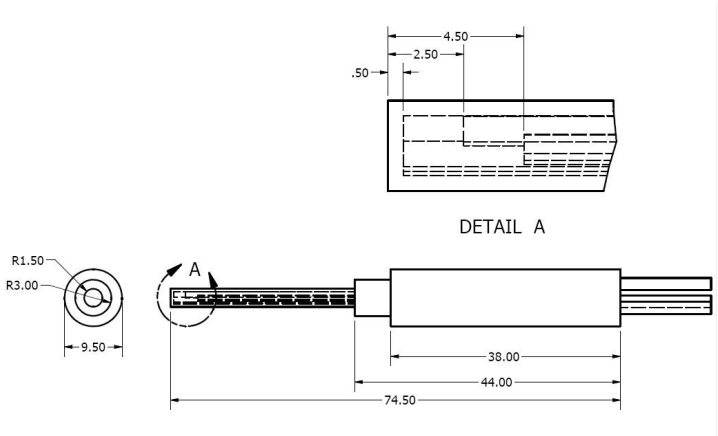


Figure 4.15: Multipoint thermocouple sensors

Sets of three thermocouples were installed in each of the five main bearing caps: the first thermocouple was used the measure oil film temperatures while the other two were used to give an estimate of the radial temperature gradient in the caps (see figure 4.16).

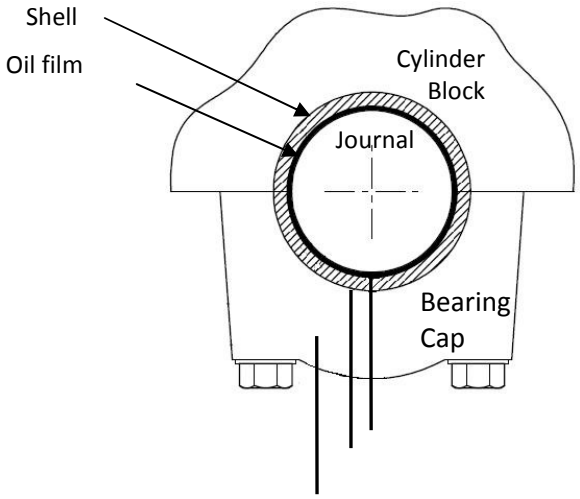


Figure 4.16: Thermocouple installation in main bearing caps

5.2.2. Data processing

Heat fluxes were estimated for different locations through the cylinder walls and from the main bearing caps. The theory applied was the same for both, only in the case of bearing caps two well-spaced temperature measurements were used whereas for the cylinder liners three measurements, separated by approximately 2mm were used. In both cases it is assumed that heat transfer only occurs radially. This is not likely to be true in areas of complex geometry such as the lower sections of the crankcase but seems a reasonable assumption in the cylindrical sections chosen here. Even so, any component of heat flux axially down the cylinder will not be apparent form the measurements. The primary aim of the instrumentation, however, is to allow differences in the thermal state of the component attributable to changes in the thermal management regime to be observed. From this perspective the assumptions made are appropriate. From first principles it can be shown that the temperature distribution radially in the cylinder liner will follow the profile described by equation 4.1 – [134]. The logarithmic term accounts for the increase in area with radius, meaning there will be a sharper temperature drop compared to equivalent heat transfer through a flat plate.

$$T_r = a \ln\left(r / r_b \right) + b \qquad \textbf{4.1}$$

Where T_r is the temperature at radius r, r_b is the bore radius and a and b are regression coefficients.

The coefficients m and b are estimated from the measured temperatures. The anticipated heat flow and temperature gradients are shown in figure 4.17. Only two measurements are required to estimate the two coefficients, however the third serves as an additional measurement for validation and as a backup in case of thermocouple failure.

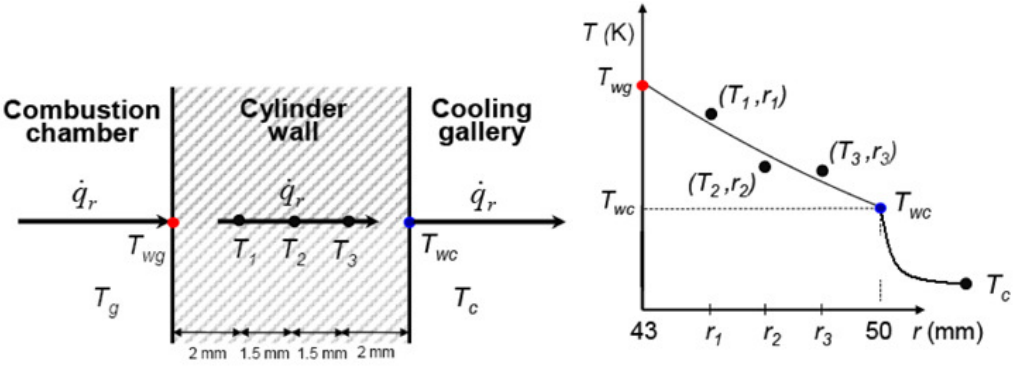


Figure 4.17: Schematic of thermal gradient estimation from temperature measurements [141]

The local heat flux normal to the cylinder liner can be derived from Fourier's equation for conductive heat transfer and is expressed for radial conduction in equation 4.2 – [141]. It is important to note that the heat flux reduces through the cylinder liner as the area increases; this explains the inclusion of the radius *r*. Throughout this work it was assumed that radial heat flow is the dominating effect. Clearly there will also be a degree of heat flow down the cylinder liner which will be more marked during warm-up and this will not be measured by the instrumentation in this study.

$$q_r = -k\frac{a}{r} \qquad \textbf{4.2}$$

where q_r is the heat flux (in W/m^2) normal to the cylinder liner at radius r (in m), k is the thermal conductivity (W/mK) of the cylinder liner and a is the regression coefficient from equation 4.1*. The negative sign indicates heat flow from higher to lower temperature.*

Using the temperature distribution estimated using equation 4.1 it is possible to estimate the temperature of the cylinder liner at the coolant side wall. Combined with the local coolant temperature measurements and calculated heat flux, applying Newton's law of cooling [134] yields an estimate for the convective heat transfer coefficient (see equation 4.3)

$$h_c = \frac{q_r}{T_{wc} - T_c} = \frac{q_r}{\Delta T} \qquad \textbf{4.3}$$

The heat transfer coefficient is an empirical term used to predict the rate of heat transfer per unit temperature difference between the cylinder wall and the coolant. As shown in chapter 2, this coefficient is affected by a number of methods:

- By increasing the coolant velocity either globally by increasing the bulk flow rate through the engine; or locally by varying the cross sectional area of the cooling passage; or by introducing an impinging jet of coolant.
- At the onset of nucleate boiling [19];
- By increasing the surface roughness of the coolant jacket.

5.3. Combustion analysis

5.3.1. Data acquisition systems

For the majority of the experiments presented in this thesis, in-cylinder pressure data was seen as essential information for assessing friction and interactions with the combustion process. The recording of in-cylinder data requires a high frequency data acquisition system, a number of fast acting sensors mounted in or in communication with the combustion chamber and a shaft encoder for alignment of measurements with crank angle position. Different combustion analysis systems have will be used, however in each case post-processing was performed or validated manually before using the resulting data:

- For steady state testing, a D2T system was used because this gave the best live feed of engine behaviour. Recording durations are limited as this system uses the base computer memory for storage and is therefore restricted. However for steady state testing, 200 cycles is largely sufficient and well within the system capabilities.
- A Kistler Kibox system was used for detailed drive cycle analysis as this system did not have any limits for data acquisition periods. Cylinder pressure from 3 cylinders, rail pressure, injection signal and fast torque could be recorded for all 17000 cycles of the NEDC at a resolution of 0.1^{o}CA. The 122M data points represent some 15GB of data per test which allows very detailed analysis of the engine behaviour. However, the analysis is time consuming and manipulating the dataset is computer intensive.
- Because the Kistler based system requires significant effort from the test cell operator during each test, for the majority of drive cycle tests a CP Engineering based system

was used which offered less performance in terms of data acquisition with many fewer channels available for continuous recording, but did allow easy and fast integration with other test results, making the data processing more reasonable for the large number of tests.

5.3.2. Engine friction measurement

Engine friction was measured using the indicator method. Brake mean effective pressure was calculated using the measured speed and torque as described in equation 4.4 - [52].

$$MEP = \frac{T \times \omega}{V_s \times \frac{\omega}{4\pi}} = \frac{P}{V_s \times \frac{\omega}{4\pi}} \qquad \textbf{4.4}$$

Net and gross indicated mean effective pressure[47] (nIMEP and gIMEP) and pumping mean effective pressure (PMEP) were obtained from the indicated results using equations 4.5 to 4.7 where the integration limits are in degrees crank angle, with 0^o representing TDC combustion.

$$nIMEP = \int_{-360}^{+360} pdV \qquad \textbf{4.5}$$

$$gIMEP = \int_{-180}^{+180} pdV \qquad \textbf{4.6}$$

$$PMEP = \int_{-360}^{-180} pdV + \int_{+180}^{+360} pdV \qquad \textbf{4.7}$$

FMEP is then calculated as the difference between nIMEP and BMEP, as described by equation 4.8. In this case, FMEP will not include pumping effects, but does include all loads from all ancillary devices. Friction, indicated or pumping torque and power may be calculated using equation 4.4. This may be integrated to estimate the respective work for a given engine duty cycle.

$$nIMEP = gIMEP + PMEP = BMEP + FMEP \qquad \textbf{4.8}$$

[47] Net indicated mean effective pressure (nIMEP) relates to the indicated work that may be used to drive the piston to overcome friction and produce useful work; it is calculated over the whole cycle. Gross indicated mean effective pressure (gIMEP) refers does not take into account pumping work and is calculated over the compression and combustion strokes only. In most engines, the pumping work is negative and relates to a loss, meaning gIMEP is higher than nIMEP, however in supercharged engines a positive pumping work can occur.

5.3.3. Heat release analysis

In-cylinder pressure signals were analysed to obtain rate of heat release (ROHR) according to a single zone heat release algorithm based on the first law of thermodynamics. Many authors have proposed more complex expressions that split the cylinder into multiple zones to allow local conditions to be predicted to assist with analysis of the emissions formation, however according to Maiboom [135] these methods are rarely validated with experimental data. The single zone analysis is the simplest and most common and assumes products and reactants are fully mixed and the gas charge is at homogeneous temperature and pressure [52]. This approach only gives an insight into the rate at which fuel energy is released, however this is sufficient for this thesis. Since the purpose of calculating ROHR here is for comparison rather than absolute accuracy it is appropriate to use a simple expression. The control volume for the single zone combustion analysis is shown in figure 4.18. The analysis involves equating the heat from combustion by calculating the indicated work on the piston, heat loss to engine coolant, mass transfers and in-cylinder temperature. The first law of thermodynamics yields the equality of equation 4.9.

$$dU_{cy} = dQ_{combustion} - dW_m - dQ_{Wall} + dm_{in} \cdot h_{in} - dm_{ex} \cdot h_{ex} + dm_f \cdot h_f - dm_{bb} \cdot h_{cy} \qquad \textbf{4.9}$$

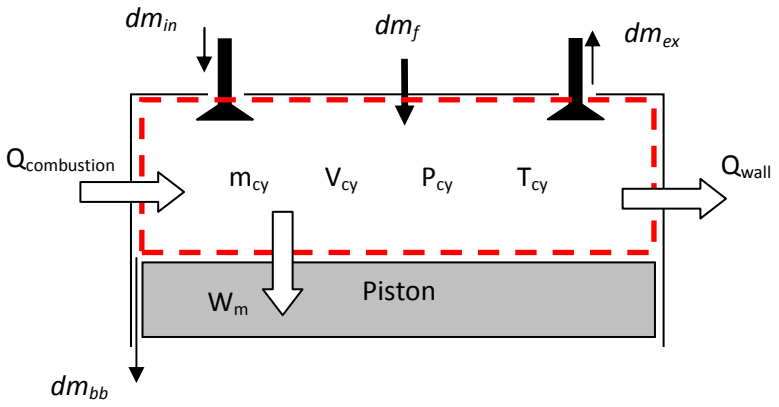


Figure 4.18: Control volume for single zone combustion analysis

The control volume considered in figure 4.18 can be simplified considerably if only the period between inlet vale closing and exhaust valve opening is considered. In this case, there is no mass flow to or from the cylinder through the inlet and exhaust valves, i.e. dm_{inlet} and $dm_{exhaust}$ are both zero. The problem can be simplified further by neglecting both blow by mass and fuel mass (see equation 4.10). Using in-cylinder pressure measurements, the bulk gas temperature

can be equated using the perfect gas law (equation 4.11) and the internal energy change follows from equation 4.12. The work can be equated as *pdV.* dQ_{wall} is obtained by an appropriate heat loss model, leaving $dQ_{combustion}$ the only unknown in equation 4.10.

$$dU_{cy} = dQ_{combustion} - dW_m - dQ_{Wall} \quad \textbf{4.10}$$

$$P_{cy} = \frac{m_{cy} \cdot R \cdot T_{cy}}{V_{cy}} \quad \textbf{4.11}$$

$$dU_{cy} = m_{cy} \cdot c_v \cdot dT_{cy} \quad \textbf{4.12}$$

When analysing pressure traces to obtain heat release data a number of issues arise where different methodologies are required and assumptions need to be made:

- Filtering the pressure signal is often necessary because heat release relies on differential of pressure which will be extremely noisy.
- Because of sensor drift the pressure data must be corrected to a known reference value at a specific point in each cycle. This procedure is refereed to colloquially as *'pegging'*.
- It is important to ensure accurate alignment of the pressure trace with cylinder volume by correctly identifying TDC on the shaft encoder as a 1^o error can result in a 10% error in IMEP [142, 143].
- Estimation of heat loss to cylinder liners is a key calculation, often poorly implemented.

Each of these aspects has received considerable attention as the impact of small errors in these aspects can have large effects on the accuracy of the heat release calculation. For the heat release analysis presented in this work, pressure signals were averaged over a minimum of 100 cycles before being filtered using a low pass Butterworth filter[48]. The absolute level of pressure was obtained using the thermodynamic method[49] [144] with reference pressure

[48] The pressure signal was filtered twice, forward and backward using the Mathworks Matlab '*filtfilt*' command which causes double the filtering effect but ensures exactly no phase lag. This is important as it ensures that the pressure signal remains aligned with the calculated volume signal and significant phase lags resulting from filtering would be equivalent to an incorrectly aligned TDC.

[49] The thermodynamic method attempts to fit a polytropic compression to the recorded pressure data using an assumed polytropic index and the pressure difference between two points. It is important the points are far enough apart to avoid noise affecting the results, but also the points must not be affected by valve closing, fuel injection or ignition.

recordings at 100° and 65° before TDC and a polytropic index of 1.37. To ensure the highest possible accuracy in the result, TDC was determined to within 0.2° crank angle according to the method by Pipitone *et al* [142]. Heat losses at the wall were calculated using Hohenberg's correlation [145] because of its simplicity.

5.4. Energy analysis

5.4.1. Instrumentation

Non-intrusive, ultrasonic coolant flow meters were used in this study to avoid impacts on pumping work. The flow meters are based on the electromagnetic flow measurement principle [146] which measures a flow velocity through a known cross sectional area. The measurement accuracy is approximately 0.3% over most of the measurement range; however, accuracy will decrease for velocities below 1m/s and measurements are not reliable for velocities below 0.1m/s. Different sized flow meters were installed to approximately match the hose sizing and anticipated flows within the circuit. For the engine-out and top hose flow meters, this represents a lower measurement limit of approximately 3l/min whereas, for the degas flow, this limit is 0.5 l/min.

Despite the flow meters themselves being non-intrusive, their installation inevitably led to additional plumbing as the devices are relatively large. This additional plumbing caused an increase in coolant volume of 2L (25%) with an associated increase in thermal inertia. It is important to account for this when comparing results to the production engine.

5.4.2. Data analysis

The majority of energy analysis in this work remains simple and is estimated based on the available measurements of flow rates and temperatures. The application of an energy balance in equation 4.13 is straight forward, where $\dot{m}$ is mass flow rate, c_p is an associated heat capacity and ΔT the temperature change. This has been applied in a number of locations in the different fluid circuits (coolant, air path and EGR gas). The specific heat capacities are well documented for the fluids involved or can be calculated from the varying gas composition in the case of EGR gases [52]. Experimentally it is important to ensure that gas thermocouples are correctly positioned to measure the stream temperatures rather than fluid film temperatures. It should also be noted that the measurement of gas temperatures with thermocouples is not

straightforward due to the thermal mass of the thermocouple and limited heat transfer between the probe and the gas. In the case of high temperature exhaust gas, conduction and radiation heat transfer from the thermocouple to the exhaust pipe can also affect the measurement [147].

$$Q = \dot{m} \cdot c_p \cdot \Delta T \qquad \textbf{4.13}$$

The energy analysis is used in this work extensively to assess heat energy flows in different heat exchangers between the fluids (EGR cooler, radiator, engine block/head, oil cooler,...). It is also used to assess thermal losses in pipes by assessing the temperature drops and treating the pipes as other heat exchangers with ambient air. When considering the system warm-up for each heat exchanger there will be an aspect of component heating and the heat lost from the hotter fluid will be less than the heat gained by the colder fluid.

6. *Chapter summary and conclusions*

The experimental facilities, candidate hardware and measurement process have been presented in this chapter. Two engines of similar age and design will be used for different objectives in this work. Prototype hardware was presented and the integration into the engine was described. The specialist instrumentation and data processing routines were also presented. The principle points to conclude this chapter are listed below:

1) Prototype hardware used in this work comprised coolant flow control valves, a dual EGR system and a variable flow oil pump. The inclusion of these devices offers an active thermal management system with a number of additional degrees of freedom.

2) The installation of the candidate hardware and instrumentation resulted in increases in water and oil circuit volumes of 2L and 0.6L respectively. This increased thermal inertia would have a negative effect on warm-up, however it was considered that in a production design these increases could be avoided.

3) The assessment of the prototype hardware performance would require a DoE approach to gain a comprehensive understanding of the system within a sensible timeframe. Each test point of the experimental design would represent a cold-start NEDC to capture dynamic behaviour.

4) The approach described in (3) provided information specific to the NEDC duty cycle. Global modelling approaches needed to be used to develop a generally applicable simulation tool.

5) Extensive instrumentation installed on the engine allowed a detailed understanding of the thermal behaviour in response to hardware changes to be measured. This was not available at all stages of the work because of development times for the instrumented engine.

Chapter 5 - *Assessing the fundamental behaviour of the thermal system*

In this chapter the fundamental behaviour of the engine in response to changes in thermal state will be presented. Experimental results both under steady state and transient conditions will be discussed.

- *In section 1, measurements of friction from hot- and cold-start NEDC will be used to show the variations in engine friction.*
- *Section 2 will then present results from steady state experiments where the thermal state is controlled by varying the engine cooling water set-point.*
- *Section 3 will assess the behaviour of varying warm-up rates using the prototype hardware described in chapter 4.*
- *Finally, section 4 will build on the drive cycle results by varying injection timing in conjunction with changes to thermal management to demonstrate potential further benefits.*

The work presented in this chapter was in part presented at the SAE world congress 2010 [148] and in part published in the proceedings of the Institute of Mechanical Engineers Part D: Journal of Automobile Engineering [149].

1. *On-engine friction measurements*

To measure engine friction using the indicator method, two drive cycles, one hot-start and the other cold start were run on the PUMA 2.0L engine whilst recording in-cylinder pressure data for multiple cylinders. Comparing the two drive cycles shows the maximum benefits achievable through faster engine warm-up. The raw calculated FMEP using this method includes significant noise and a filtering process is required[50]. It is important before analysing the filtered data to assess the impact of the filter itself against the raw data (see figure 5.1). The filter clearly allows a better assessment of the steady state periods; however it is important to note the effect on the transient events which are slightly distorted. Despite this, it serves as a good tool for assessing differences as the engine warms up over the drive cycle.

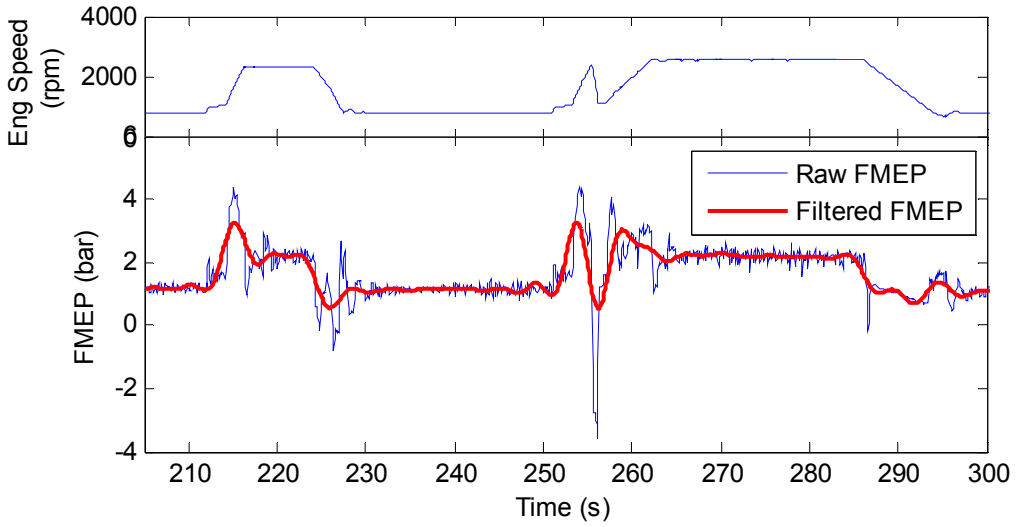


Figure 5.1: Impact of filter on FMEP measurement over section of NEDC

Figure 5.2 shows filtered FMEP for each repeat of part of the UDC cycle over the hot and cold-start tests. Here the four repeats of a section of the ECE cycle as the test progresses from cold have been superimposed, along with the same data from the fully warm cycle, so as to allow comparison. The gradual reduction in FMEP over the cold-start drive cycle shows the temperature influence on friction. Results from the hot test represent a further reduction in friction and overlay each other showing good repeatable levels throughout the test. This shows that as the engine warms up from a start oil temperature of 25°C to the fully warmed up condition (approximately 100°C in the sump) friction levels drop by more than half, from an

[50] A low pass Butterworth filter was used using the Mathworks Matlab *'filtfilt'* - double filtering approach to avoid phase lags.

average FMEP of 2.7bar to 1.3 bar at an engine speed of 2500rpm. It is also of interest to note the influence of engine speed/load operating point on friction which is of the same order of magnitude as that of temperature.

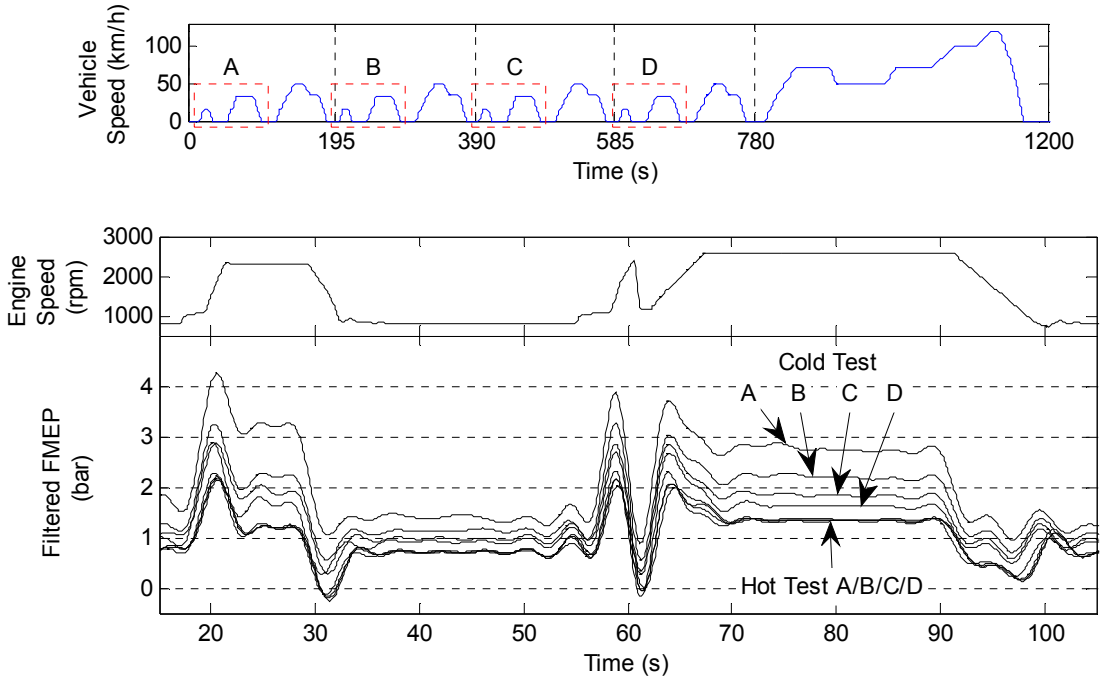
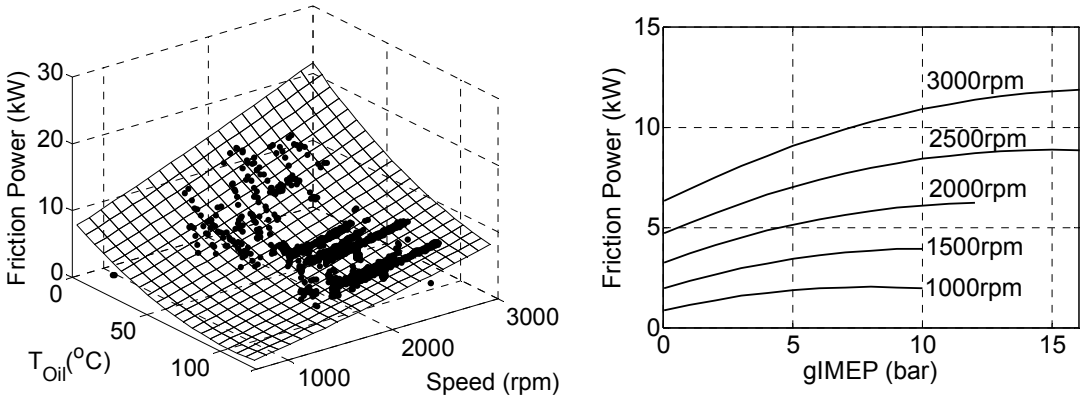


Figure 5.2: FMEP over part of UDC for cold-start test and hot-start test showing reduction in friction during warm-up – The upper frame shows the entire NEDC, sections marked A to D are the initial section of the ECE cycle, which itself is repeated four times to comprise the ECE15 or urban phase of the NEDC. The second frame shows engine speed and filtered FMEP, enlarged for sections A to D. The resulting plot shows the reduction in friction during warm-up.

Because of the multiple dependencies of friction on speed, load and temperature it is difficult to visualise the data on an individual scatter plot and a polynomial model[51] has been fitted to the data. The model has three inputs, Speed, gross IMEP and oil temperature and predicts friction power[52] (kW). Figure 5.3(a) shows a wire frame surface plot representing the response model prediction of friction power as a function of engine speed and oil temperature for a constant gross IMEP of 10bar. Also shown as a scatter plot, superimposed on the surface are the raw data points for which gIMEP lies between 8 and 12bar.

[51] A quadratic model with 1st order interaction terms gave a good fit with an R^2 and PRESS R^2 of 0.85 which was very good considering the noise in the measured data.

[52] When fitting the model the raw friction data may be used rather than the filtered data as the averaging aspect of the simple regression model will account for this aspect.



(a) *(b)*

Figure 5.3: Friction power response model for (a) engine speed and oil temperature (gIMEP 10bar Raw data for gIMEP between 8-12bar also shown) and (b) gross IMEP (curves truncated to reflect engine operating space)

The response model confirms that engine speed and oil temperature each have an effect on friction of a similar magnitude as they vary over the range observed in the drive cycle. The model suggested a smaller increase in friction due to increased IMEP (figure 5.3(b)). This effect will be the result of two opposing phenomenon: as load increases gas pressures will increase loading of piston rings, piston skirt (due to an increase in the radial component of gas load reacted by the connecting rod when the connecting rod is at an angle to the cylinder axis), valve train mechanism, main and big end bearings, whereas the higher in-cylinder temperatures will cause locally higher oil temperatures reducing hydrodynamic friction. The results suggest the latter is the dominant effect resulting in an impact approximately 40% that of speed and oil temperature. The results obtained here from dynamic measurements are similar to the results presented by Piptone [150] from steady state measurements and empirical friction models used by other authors [39, 40].

These results are analogous to the Stribeck curve presented in chapter 2:

- Higher engine speed translates into higher sliding velocities at the contact surfaces. This behaviour is typical of hydrodynamic operating regime.
- Higher oil temperatures equates to lower lubricant viscosities. This is also typical of hydrodynamic friction regime.
- Higher IMEP equates to higher pressures and will affect boundary friction in the valve train and piston rings.

The friction measurements may be integrated over the NEDC to give an estimate of total friction work. This is shown in figure 5.4 for both the hot and cold-start drive cycle. The brake work and pumping work are similar in both tests, representing approximately 8.3MJ and 1.3MJ respectively. This is due to the test being conducted on a dynamic engine dynamometer rather than in vehicle. In a vehicle the tyre losses would be significantly greater when cold, leading to an increase in brake work over the cycle. Typically a cold cycle on chassis dynamometer will use 8% more fuel that a warm cycle [36].

The friction work is different with 5.3MJ and 3.7MJ of energy used to overcome friction in the cold-start and hot-start tests respectively. In the cold-start test, this represents 36% of total indicated work. Over the drive cycle, a 30% reduction in friction is observed by virtue of starting with a hot engine. This represents an energy saving of 1.6MJ which, based on the lower heat value of the fuel, is equivalent to 3% of the total fuel energy usage, reflecting the measured differences in fuel consumption from the two tests.

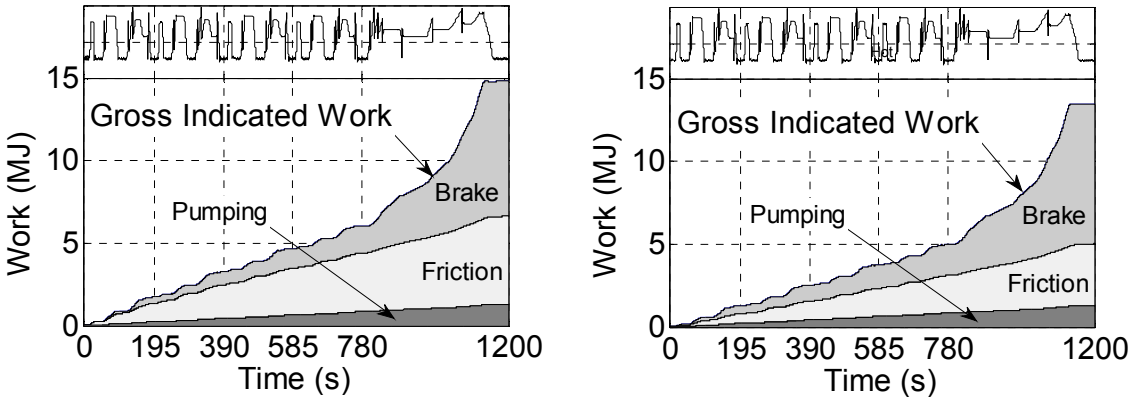


Figure 5.4: Mechanical energy usage of indicated work for (a) cold-start and (b) hot-start NEDC

2. *Variation of engine thermal state*

Steady state tests were performed on the PUMA 2.0L engine with the aim of assessing the effect of varying engine coolant temperature on both the physical behaviour and the control strategy of the engine.

2.1. Test Plan

All tests were run under steady state conditions with measurements representing average values over a 30 second period. The analysis was repeated for two operating points taken from the NEDC: the first, representing the 50km/h condition was an engine speed of 2424rpm and a torque of 26.4Nm and the latter representing 100km/h was set at 2226rpm, 144Nm. In all cases the operating point was kept constant and the host system adapted the pedal activation to maintain constant brake output. As a result, BMEP remained constant and analysis of the change of friction with temperature was undertaken by observing changes in IMEP.

A design of experiments approach was adopted to better explore the operating range of the engine over temperature, EGR and injection timing (SOI) settings and table 5.1 details the experimental design:

- Coolant temperature set-points were[53]: 50^{o}C, 70^{o}C, 86^{o}C and 98^{o}C.
- Injection timing set-points were the production engine setting, 2^{o}CA advance and a 2^{o}CA retard to remain within sensible operating regime with respect to emissions calibration.
- EGR set-points were again set relative to the production setting by increasing or reducing the EGR valve[54] opening by 1mm. It is important to note also at this point that although the changes in EGR valve setting were linear, the changes in EGR rate were highly non-linear.

[53] 86^{o}C was considered as the production engine operating point while 98^{o}C was estimated as an upper limit to avoid excessively high metal temperatures. The lower set points of 70^{o}C and 50^{o}C were chosen to cover a wide operating range.

[54] As this has a significant effect on boost pressure, this was adjusted with the aim of keeping boost pressure constant by changing turbocharger (VGT) settings, although in some cases the VGT became chocked and constant boost was not achievable.

When generating the response models, the true measured points were taken rather than the desired settings. This also included using EGR rate, based on the ratio of CO_2 concentration, rather than EGR valve position. It was also noted that whilst the aim was to achieve steady state operation, due to the large thermal inertia of the engine and cooling system, very low frequency variation of engine temperature occurred during testing however variations during any single measurement were less than 3°C.

#	Coolant temperature	SOI	EGR rate	#	Coolant temperature	SOI	EGR rate
1	50°C	2° Adv	Low	19	86°C	2° Adv	Low
2			Standard	20			Standard
3			High	21			High
4		Standard	Low	22		Standard	Low
5			Standard	23			Standard
6			High	24			High
7		2° Rtd	Low	25		2° Rtd	Low
8			Standard	26			Standard
9			High	27			High
10	70°C	2° Adv	Low	28	98°C	2° Adv	Low
11			Standard	29			Standard
12			High	30			High
13		Standard	Low	31		Standard	Low
14			Standard	32			Standard
15			High	33			High
16		2° Rtd	Low	34		2° Rtd	Low
17			Standard	35			Standard
18			High	36			High

Table 5.1: Experimental design for steady state testing

2.2. Engine strategy analysis

For EGR and SOI, the tests were run at points relative to the production calibration. As a result, certain elements of engine operating strategy were revealed and these compromised the test plan. Figure 5.5 shows the actual testing points of the experimental design with respect to the three inputs.

Figure 5.5 (a) shows the interactions between coolant temperature and injection timing. There is an offset in injection timing between tests taken below 70°C and those above 86°C. At the cooler operating temperatures, injection timing is advanced by approximately 3°CA which will allow for better combustion during colder conditions. If injection is retarded significantly at lower operating conditions, there will be a larger risk of misfiring as gas temperature in the cylinder may not be sufficiently high. In contrast, at higher operating temperatures, injection will be retarded to limit peak in-cylinder temperatures for NO_x control. At the lower temperatures, NO_x formation will be less problematic meaning injection may be advanced

without detrimental effect. The results presented here are for the 50km/h case, but similar observations were made at the higher load condition.

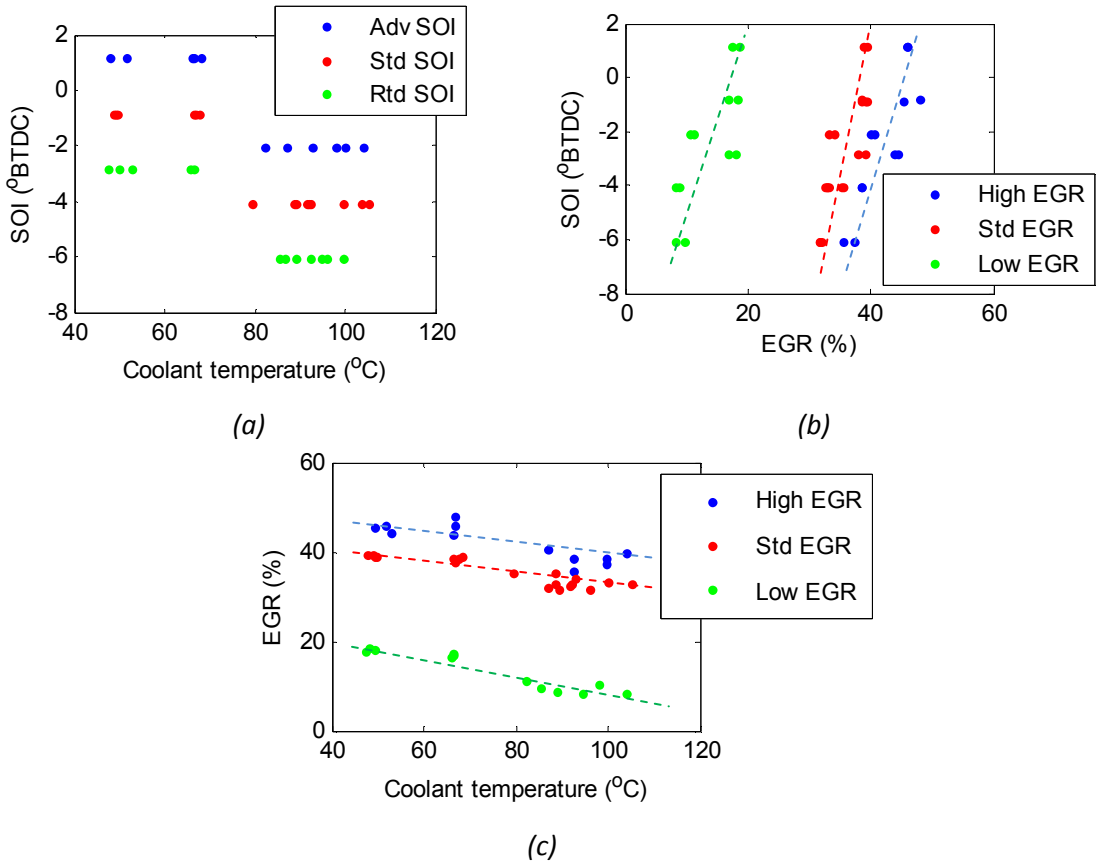


Figure 5.5: Actual experimental design for (a) injection timing and coolant temperature, (b) injection timing and EGR rate and (c) EGR and coolant temperature

Figure 5.5 (b) shows operating points with respect to EGR and injection timing. There is very little correlation between the two variables; however there is a significant non-linearity in the distribution of EGR rate as the step between the low setting and standard setting is much larger than between the standard and high settings. This is because EGR rate was controlled by the opening of EGR valve and the relationship between EGR valve opening and EGR rate is highly non-linear.

Finally, figure 5.5 (c) shows interactions between coolant temperature and EGR rate. In this case there is a gradual, almost linear change which suggests the interaction may be more physical rather than due to the engine calibration. This may be caused by the EGR gas cooler, which relies on coolant to evacuate heat. At lower operating temperatures there will be more potential for cooling EGR gases which will increase the density of these gases and increase the EGR rate.

2.3. Fuel Consumption and NO_x emissions

The measured NO_x and fuel consumption were used to calculate *fuel consumption improvement per unit NO_x penalty* and *NO_x improvement per unit fuel consumption penalty*. This allows the different controls to be compared equally, despite the larger scope for variations using EGR and SOI [148]. Table 5.2 details these rates where potential improvements are obtained from local gradients in the response model. In each case, the quoted improvements are relative to the "standard" operating points in terms of the three parameters[55].

	Operating Point 1		Operating Point 2	
Parameter	% bsfc/ %NO_x*	% NO_x/ %bsfc#	% bsfc/ %NO_x*	% NO_x/ %bsfc#
Injection Timing	0.14	4.0	0.07	6.7
Coolant temperature	0.10	10.0	0.10	20.6
EGR rate	0.01	13.0	0.04	9.7

* refers to the improvement in bsfc that is obtained for 1% deterioration in NO_x

\# refers to the improvement in NO_x that is obtained for 1% deterioration in bsfc

Table 5.2: Comparison of NO_x and fuel consumption improvement potential for injection timing, coolant temperature and EGR rate

From the results it is immediately obvious that the coolant temperature offers similar, if not better fuel consumption and NO_x trade-offs than adjustments to injection timing or EGR rate. This is surprising as engine thermal management has rarely, if ever, been used in the past on production vehicles to control emissions. This is, however, consistent with the work reviewed in chapter 2 [23]. This is probably best explained by the difficulties and smaller scope for improvements through variations in the thermal system however, this may become essential to meet future design requirements.

2.4. Combustion effects

Figure 5.6 shows in-cylinder pressure traces for varying coolant temperatures at both operating points. Figure 5.7 shows the respective rates of heat release. Concentrating firstly on results from operating point 1, it is obvious that there are two groups due to the shift in SOI

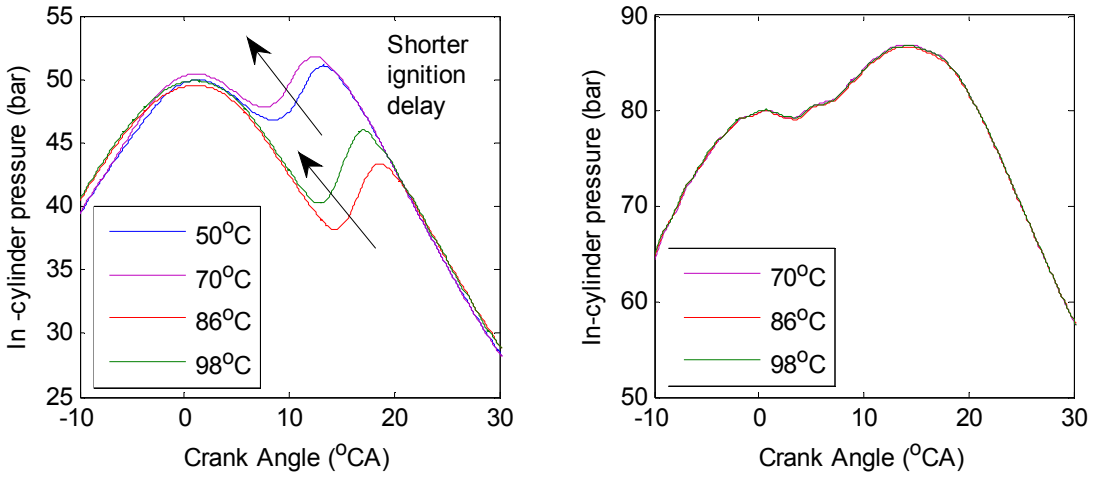
[55] For example, the quoted improvement in bsfc per NO_x penalty for coolant temperature at operating point 1 that is quoted as 0.1% bsfc per percentage NO_x is obtained by increasing coolant temperature. Conversely, the improvement NO_x per fuel consumption penalty is obtained by reducing coolant temperature.

(see figure 5.5 (a)). If the two groups of curves (50°C / 70°C and 86°C / 98°C) are treated separately, then the effects of coolant temperature become apparent:

- Higher coolant temperatures reduce ignition delay.
- The shorter ignition delay reduces the pre-mix combustion spike.

The shorter ignition delay is apparent both in the cylinder pressure and heat release traces. At hotter coolant temperatures, the pressure rise due to combustion occurs sooner in the cycle despite identical injection timing. Similarly, heat release also occurs sooner in the cycle. The hotter conditions will be more favorable for fuel evaporation and allowing the fuel faster mixing and reaching its self-ignition temperature sooner. The fuel itself may also be hotter when injected into the cylinder. The reduced premix combustion is a direct consequence of shorter ignition delay as less fuel will contribute to this phase (see chapter 2).

If all else remains constant, increased ignition delay and premix spike can cause an increase in NO_x emissions as premix combustion is responsible for the rapid increase in local flame temperature. However, NO_x emissions were seen to decrease with lower coolant temperature which suggests that overall engine temperatures are the dominant effect in this process.



(a) *(b)*

Figure 5.6: Cylinder pressure for different target coolant temperatures at (a) operating point 1 and (b) operating point 2. At operating point 1, SOI and EGR rate are identical for 50°C/70°C and for 86°C/98°C. At operating point 2, SOI and EGR are identical for all coolant temperature set-points.

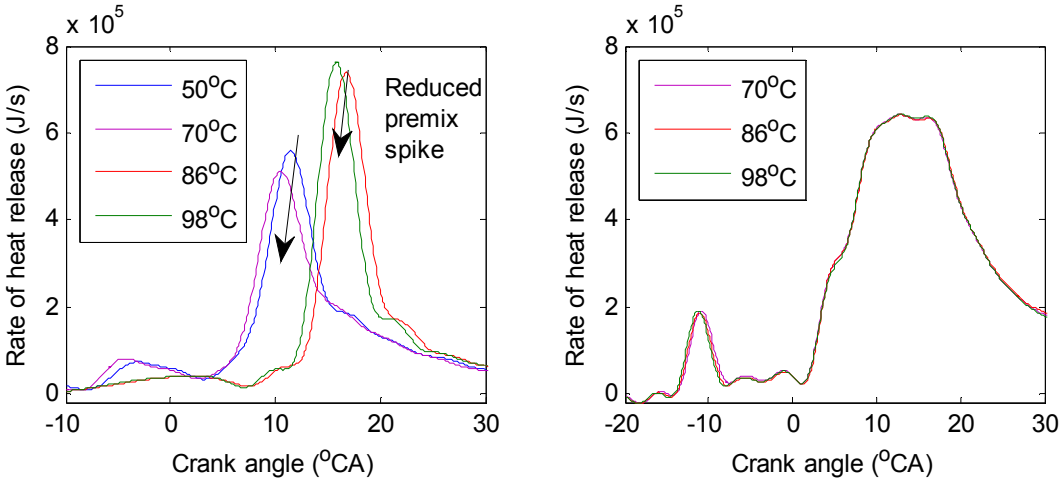


Figure 5.7: Rate of heat release for different target coolant temperatures at (a) operating point 1 and (b) operating point 2

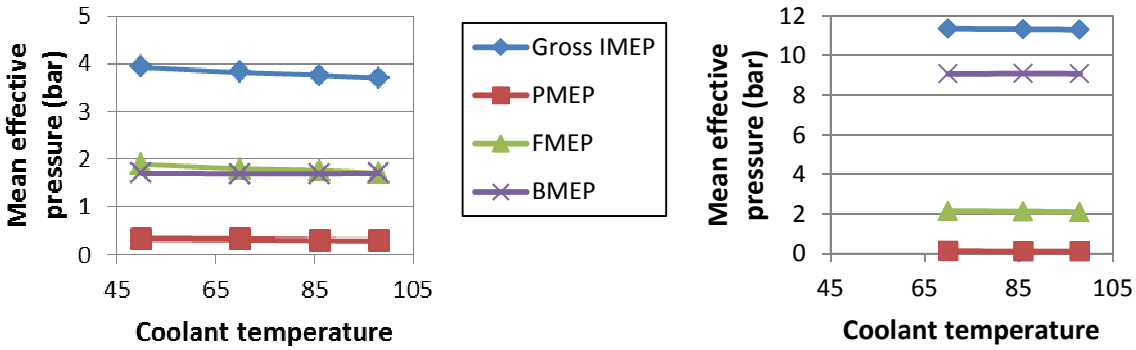
When looking at cylinder pressure and ROHR from the higher load point (operating point 2) it is obvious that there is only one group of results[56]. Secondly, there are no significant differences between pressure or heat release. The most significant differences can be seen around the pilot injection, but these remain small. The reduced effects of coolant temperature on the combustion process may be a result of higher local coolant temperatures around the combustion chamber with the possible onset of nucleate boiling that could render metal temperatures essentially insensitive to bulk coolant temperatures however, These discrepancies between load points are consistent with the work by Torregrosa *et al.* [89].

Mean effective pressures across the range of coolant temperatures at production calibration SOI and EGR rates are summarized in table 5.3 and plotted in figure 5.8. Each value quoted is the average of two independent repeated measurements, each comprising of the average of 100 cycles. The results are also averaged over the 3 instrumented cylinders. At operating point 1, FMEP tends to reduce by 0.2bar (10%) as the coolant temperature is increased from 50°C to 98°C. At operating point 2 the reduction in FMEP is much smaller and about 0.05bar (2.5%) when coolant temperature is raised from 70°C to 98°C. This suggests that friction is less sensitive to temperature at higher loads, but these results could also be explained by measurement difficulties due to the relative contribution of friction to overall indicated work at each operating point.

[56] Because the test run with a target coolant temperature of 70°C were actually run with temperatures higher than 70°C (rather than slightly below the target temperature as for operating point 1) the two groups in this case were 50°C and 70°C / 86°C / 98°C. As a result, the 50°C coolant temperature case has been omitted from the results.

	Operating Point 1			Operating Point 2		
Coolant temperature	**Gross IMEP**	**PMEP**	**FMEP**	**Gross IMEP**	**PMEP**	**FMEP**
50	3.93	0.34	1.89	N/A	N/A	N/A
70	3.82	0.33	1.80	11.35	0.13	2.15
86	3.76	0.31	1.76	11.33	0.11	2.13
98	3.70	0.30	1.70	11.30	0.12	2.10

Table 5.3: Mean effective pressures for temperature swings for both operating points



(a) *(b)*

Figure 5.8: Variations of mean effective pressures at different coolant temperatures for (a) operating point 1 and (b) operating point 2

The changes in FMEP are considerably smaller than those measured during warm-up, presented earlier in this chapter. This highlights the difference between *an engine during warm-up* and *a hot engine with cold coolant*. The results presented by Wakuri et al [54] and Daniels and Braun [55] also differ significantly from these results, despite also using controlled coolant temperature. These differences could be a result of the different thermal soaks in the respective testing campaigns, however there is insufficient information comment further.

3. *Variations to engine warm-up rate*

3.1. Test plan

The alternative cooling system was installed on the non-instrumented 2.4L PUMA engine. A number of calibrations were tested over the NEDC cycle using a combination of the new actuators[57]. Each experiment was a cold-start NEDC and these are listed in table 5.4. Four of these setups have been identified for further analysis and identified by a *designation* in the list of calibrations.

[57] Although the VFOP is included in this engine build for practical reasons, analysis of the effect of varying oil flows will be minimal in this chapter and will be covered in detail in chapter 6.

Designation	Engine-out coolant throttle (opening temperature[#])	EGR cooler coolant flow (% of max flow at idle[&])	Oil cooler coolant flow	EGR cooler	VFOP
Uncontrolled	Fully open	Non Controlled	Open	Coolant	Max Flow
Baseline	Fully open	Non Controlled	Open	Coolant	Controlled*
Build 2	Mapped (95°C)	Non Controlled	Open	Oil	Controlled*
Build 3	Mapped (105°C)	Restricted (8%)	Open	Oil	Controlled*
N/A	Mapped (90°C)	Max	Open	Coolant	Controlled*
N/A	Mapped (90°C)	Restricted (50%)	Open	Coolant	Controlled*
N/A	Mapped (90°C)	Max	Bypassed	Oil	Controlled*
N/A	Mapped (90°C)	Restricted (17%)	Open	Oil	Controlled*
N/A	Mapped (90°C)	Restricted (17%)	Bypassed	Oil	Controlled*
N/A	Mapped (90°C)	Max	Open	Oil	Controlled*
N/A	Fully open	Max	Open	Oil	Controlled*
N/A	Mapped (108°C)	Max	Open	Oil	Controlled*
N/A	Mapped (105°C)	Max	Open	Oil	Controlled*
N/A	Mapped (100°C)	Max	Open	Oil	Controlled*

* Controlled oil flow is based on a target oil gallery pressure of 2bar

Opening temperature represents the cylinder head metal temperature (the standard ECU engine temperature measure for this design of control system) below which engine-out coolant throttle remains closed. Above this temperature the throttle will gradually open.

& EGR cooler coolant throttle was set to restrict flow in the heating loop. The opening was set according the measured flow at idle, which was measured as a percentage of the flow when the valve was fully open.

Table 5.4: List of all experimental setups considered during scoping exercise

3.2. Engine Thermal behaviour

The effect of the different control strategies during warm-up on coolant temperatures and energy exchanges is shown in figure 5.9. Analysis of the engine-out coolant temperature clearly shows faster warm-up as the coolant flows are progressively throttled through builds 2 and 3. The key differences are seen in the main coolant loop around the radiator and degas bottle. Measurements from the flow meters suggest that there is no flow from the engine to the thermostat, however the temperature measurements suggest there is a small leakage flow. The effect of closing the engine-out throttle is clearly seen at the top hose where there is no flow until the coolant throttle opens. Detailed analysis of the coolant flows shows there is minimal flow through the radiator in all builds, but that in the baseline setup there is a significant flow through the Degas bottle. Although the main engine-out throttle has not completely isolated the front end coolant circuit, it has isolated the degas bottle. No measurements were taken to explain definitively why this was the case, but it is believed to be a result of changing pressure distribution in the circuit[58]. A significant benefit from this will be to reduce the thermal inertia of the coolant system during warm-up. The throttle was located in a similar position to other studies [34, 38, 89] and thermal inertia was reduced by the same principle as these studies.

[58] Coolant pressures were measured only at pump inlet and outlet and these do not show significant differences with respect to engine throttling.

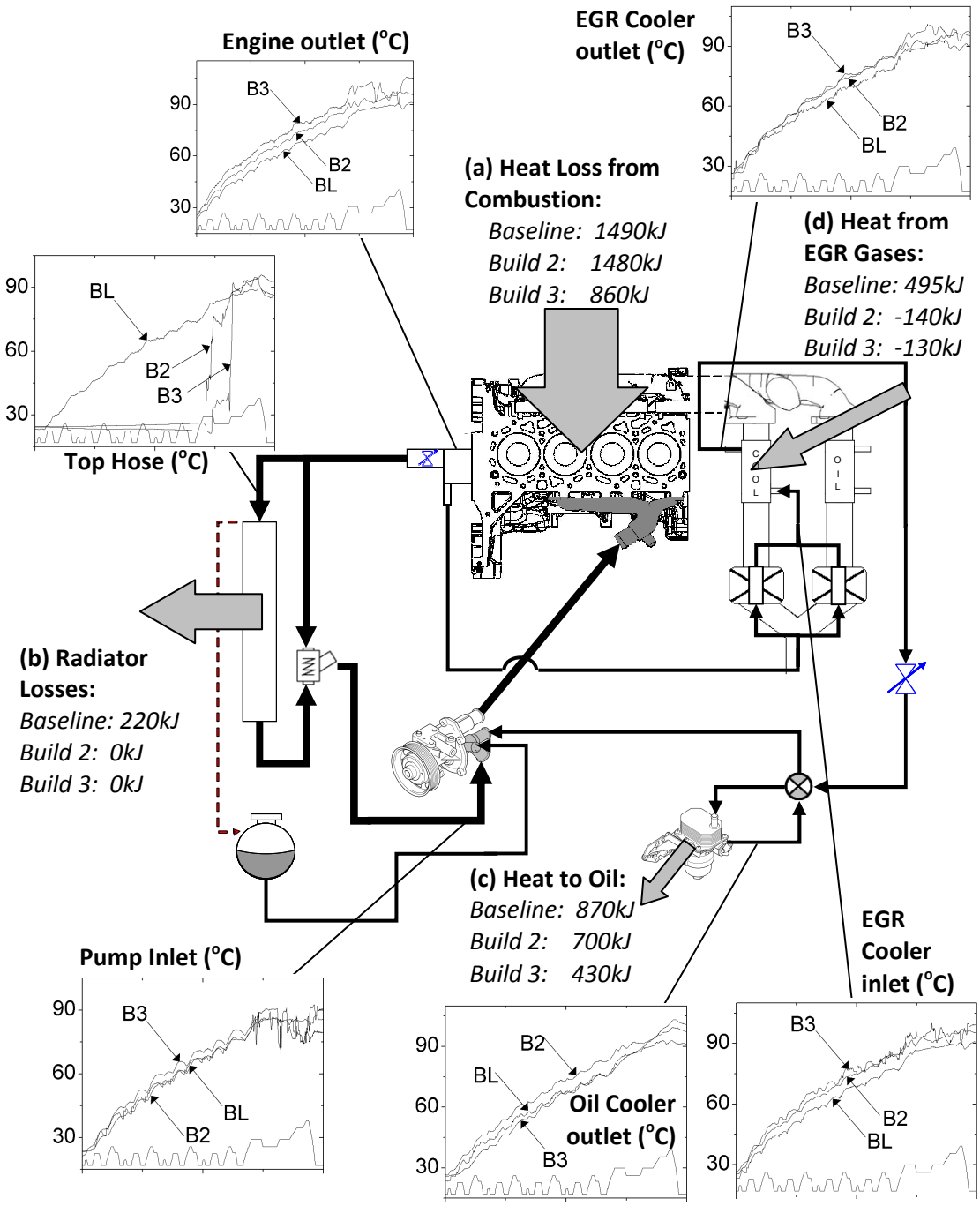


Figure 5.9: Energy exchanges and selected coolant temperatures for baseline, build 2 and 3 (all temperatures in °C, all plots against time over NEDC with vehicle speed superimposed for reference)

The effect of throttling coolant flow through the oil cooler reduces the heat transfer to oil in this device. This is seen by the lower calculated heat loss from the coolant (see figure 5.9) and the smaller rise in oil temperature over the external circuit in figure 5.10. Also shown is the oil temperature change for the production engine over its external circuit which does not include the additional hardware. This shows that the additional heat input to the oil is sufficient to overcome the additional thermal inertia and ambient losses from the exposed circuit.

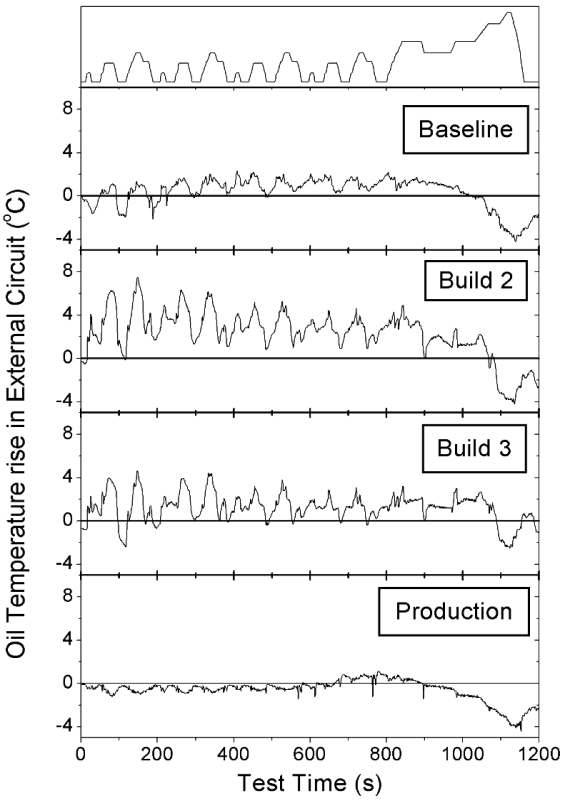


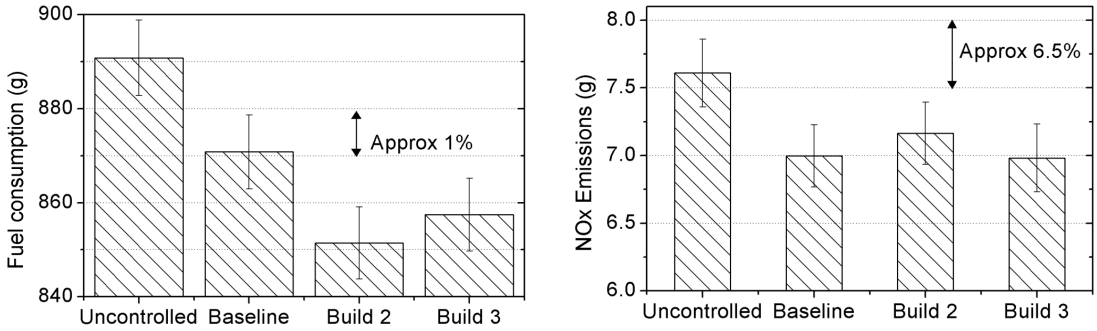
Figure 5.10: Oil temperature difference over external oil circuit for baseline, build 2, build 3 and production engine (without oil EGR cooler hardware)

Maintaining the main engine-out coolant throttle closed over the warm-up period effectively removed the volume of coolant in the radiator and degas bottle from participating in the warm-up, effectively reducing the thermal inertia. In addition, the reduced heat transfer to the coolant suggests more heat remains in the structure. It is important to note that it is the combination of the reduced thermal inertia of the coolant and the reduced heat transfer that allow the engine to warm-up faster. Lower heat transfer to a similar mass of coolant would simply allow the structure to run at a higher temperature compared to the coolant, but the coolant warm-up rate would be slowed down. Reducing the coolant inertia allows it to warm-up faster, whilst also allowing the structure to run with a higher temperature differential.

The dual EGR cooler and oil cooler control does not significantly change the available heat or inertia. Instead this system can modulate heat flows around the engine. This was seen by changes in oil temperature across the external circuit. Because of the limited instrumentation, it is unclear whether this increased temperature results in hotter oil at key friction locations in the engine. To reduce viscous losses, it is necessary to have hotter oil at areas of hydrodynamic lubrication. It will be important to investigate how well the oil retains the additional heat supplied to it or whether this is lost very quickly to the cold structure.

3.3. Fuel consumption and NO_x analysis

Fuel consumption for each of the four setups is shown in figure 5.11 (a) along with 95% confidence intervals based on the engine/dynamometer facility performance established in chapter 3. The first point to note is that changes in fuel consumption are small and of the order of 1%. The largest change is between uncontrolled and baseline tests which are accounted for by the VFOP (22g or 2%); this will be analysed further in chapter 6. Heating the oil using EGR gases and throttling coolant flow during warm-up yielded a further 20g (2%) improvement when considering the gravimetric measure. However, throttling the coolant flow further led to a slight increase in fuel consumption (approximately 5g) which will be shown to be caused by interactions with the engine control system. NO_x emissions are detailed in figure 5.11 (b) again with 95% confidence intervals[59]. These results follow the opposite trend to FC results.



(a) *(b)*

Figure 5.11: (a) NEDC fuel consumption and (b) NO_x emissions from different engine calibration settings

Fuel consumption changes between the *uncontrolled* and *baseline* calibrations will be discussed in chapter 6. Faster engine warm up in *build 2* calibration relative to the *baseline* appears to result in reduced frictional losses and lower fuel consumption. Restricting flow further in the case of *build 3* improved coolant water warm-up further, but this did not result in further fuel consumption benefits.

Other emissions were also affected by the changes in warm-up rate. These results are shown in table 5.5 for HC and CO. Both are reduced in *builds 2 and 3* compared to the *baseline* which is consistent with results from other studies in the literature [26, 28, 34].

[59] Although the production engine was EURO IV specification, these results exceed those limits because of a fault in the EGR strategy due to test bed/engine interactions. This fault was corrected for subsequent experimentation, but affects all results in this chapter. The fundamental effects of the thermal management changes are unaltered.

Test	HC	CO
Baseline	4.7g	23g
Build 2	*-6.8%*	*-7.6%*
Build 3	*-6.2%*	*-7.4%*

Table 5.5: HC and CO emissions from baseline test and percentage improvement in subsequent builds

When considering the physical mechanisms, the change in FC and emissions that occurs between the *baseline* and *build 2* calibrations is not surprising. Higher engine temperatures sooner in the cycle would reduce engine friction, improve combustion and increase the reactions described by the Zeldovitch. The results from *build 3* calibration are less expected. Even though the engine appears to run hotter than *build 2* (see coolant temperatures, figure 5.9) NO_x emissions decrease while FC, CO and HC all increase slightly.

To better understand these results, the NO_x formation throughout the NEDC is shown in figure 5.12. The figure shows cumulative NO_x and the difference in cumulative NO_x relative to the baseline setup. Firstly it is obvious that the final NO_x result is heavily dependent on the final 150 seconds: approximately 75% of NO_x are formed during the 100kph and 120kph cruises and the associated transients. The particularity of these last high power points of the cycle is also apparent when considering the difference graphs: for *builds 2 and 3*. Less NO_x are produced over the first part of the drive cycle, however over the final 150 seconds significantly more NO_x is produced giving a net increase in the case of *build 2* and similar overall NO_x in *build 3*. A further understanding of this phenomenon will be provided by analysis into the ECU behaviour.

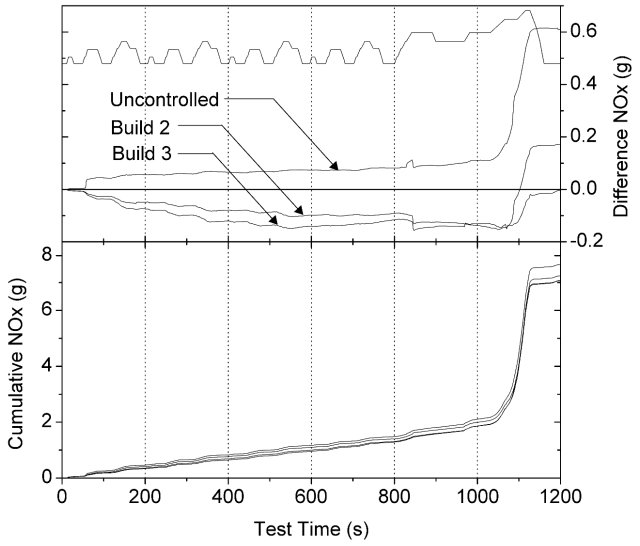
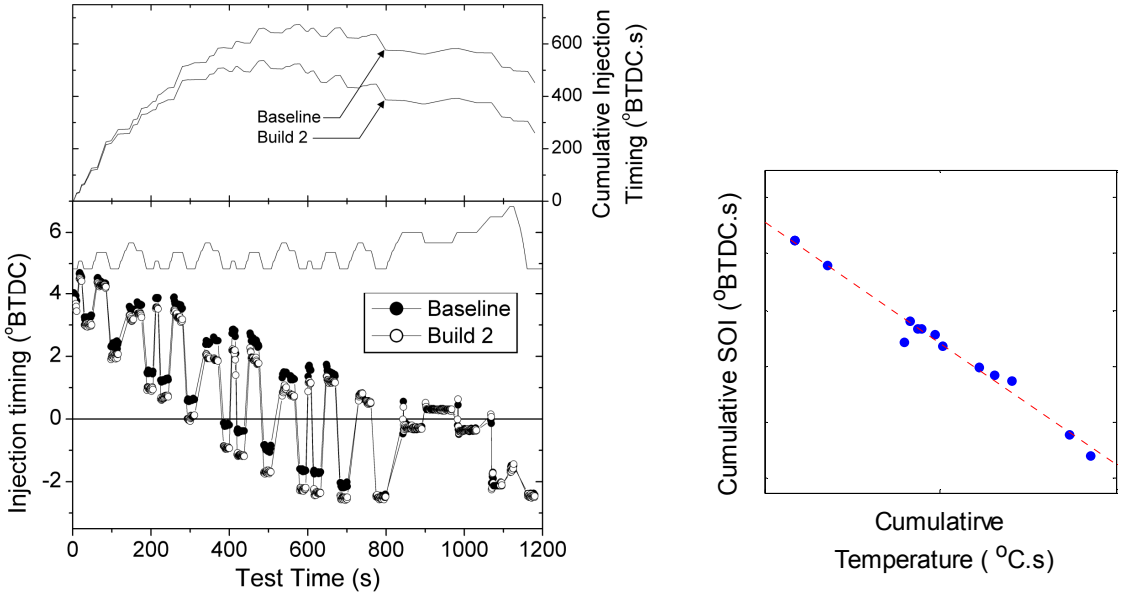


Figure 5.12: Cumulative NO_x and difference-to-baseline for each test over NEDC

3.4. Engine strategy analysis

The steady state testing in section 2 of this chapter highlighted strong interactions between temperature and key control factors such as injection timing and EGR. Figure 5.13 (a) shows the injection timing for *baseline* and *build 2* calibrations. Also included is the cumulative injection timing which has little physical meaning, but is a useful tool for showing small differences between the transient signals[60]. In both tests SOI is gradually retarded as the engine warms up. In *Build 2* the ECU clearly retards sooner as the cumulative graphs diverge over the first 600 seconds due to the difference in warm-up rates. When the engine temperature sensor exceeds about 80°C the cumulative graphs stop diverging and injection timing is the same for both tests. Figure 5.13 (b) shows the correlation between injection timing and engine warm-up by plotting cumulative injection timing against cumulative cylinder head temperature. These results are obtained from all calibrations detailed in table 5.4. Faster warm-up rates result in higher value of cumulative temperature which correlate with lower values of cumulative SOI, i.e. more retarded injection timing.



(a) *(b)*

Figure 5.13: Interactions between injection timing and engine temperature: (a) Injection timing comparison for baseline and build 2 where faster warm-up causes retarded injection earlier in the cycle; and (b) Correlation between cumulative SOI and cumulative engine temperature over urban drive cycle for all calibrations, again showing retarded injection timing with faster warm-up.

[60] Injection timing is expressed in degrees before top dead centre (°BTDC), meaning a lower or negative value indicates a more retarded condition. Cumulative injection timing is similar in as much as a lower value indicates a tendency to run more retarded.

EGR rates have been studies in a similar way. The measurement of EGR rate relies on two measurements of CO_2 concentrations and although care has been taken to correctly time align these measurements, small differences during transient events mean a reliable measure is not possible. For display purposes filtering has been applied[61]. Figure 5.14 (a) compares the raw and filtered EGR rates to show the signal retains integrity. Figure 5.14 (b) shows the EGR rate for the *baseline* and *build 2 and 3* calibrations. Although some small differences exist over the first 200 seconds, these appear to be mostly during idle phases and for the majority of the cycle, EGR rates are similar. During the EUDC phase the previously mentioned closing of the EGR valve due to engine/test facility interactions is apparent[62].

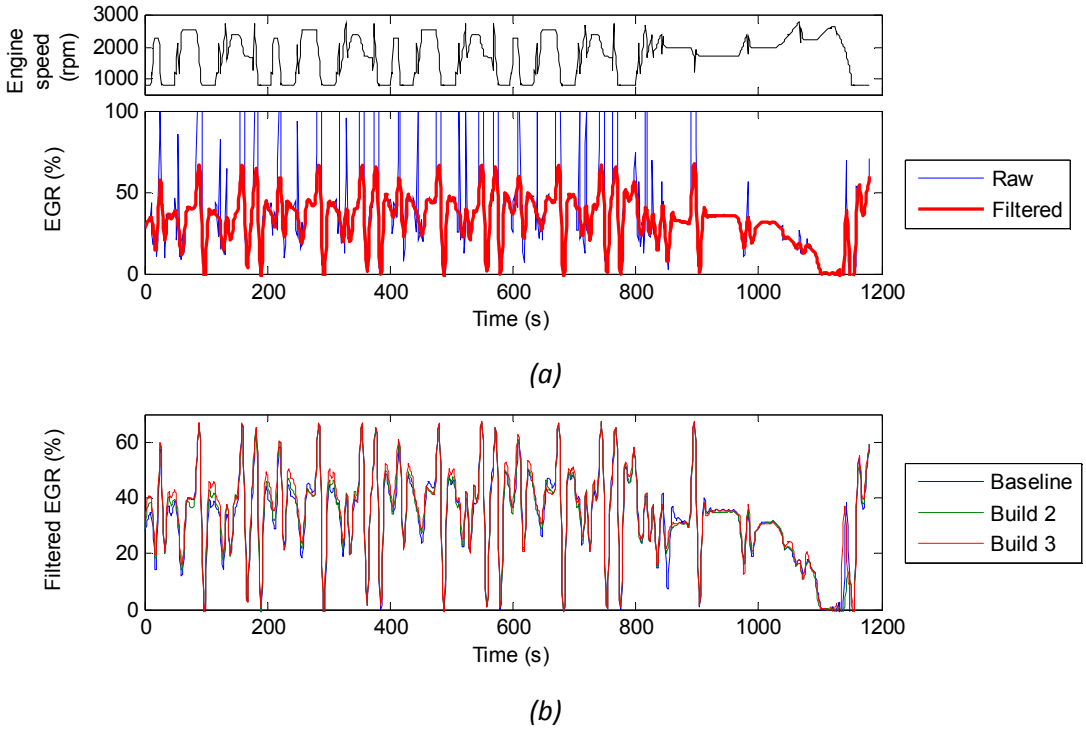


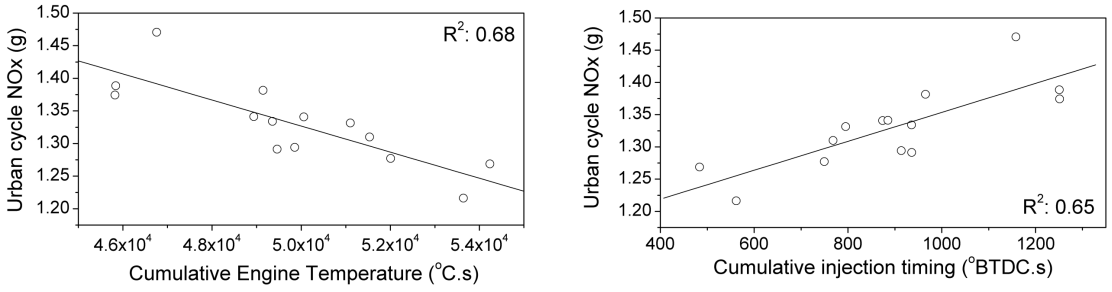
Figure 5.14: (a) Effect of EGR filtering over NEDC; (b) Filtered EGR rate for baseline and builds 2 and 3; EGR rate based on ratio of inlet and exhaust CO_2 measurement.

The change in NO_x over the first 600 seconds in *builds 2 and 3* compared to the *baseline* calibration will be the net effect of increased engine temperature and retarded injection timing. Referring back to the NO_x results in figure 5.12, the effect of injection timing would appear larger as there is an overall reduction in NO_x. This suggests an over-compensation by the engine ECU due to the increase in temperature which will be compromising the potential benefits in fuel consumption. This is caused by the engine temperature being estimated based

[61] A combination of signal clipping and subsequent low pass Butterworth filter was used
[62] This was corrected in subsequent work through modifications to the engine strategy.

on a single temperature measurement in the cylinder head that is not representative during engine warm-up. Over the latter part of the NEDC when injection timing is similar in all tests, the changes in NO_x will be a result of differences in thermal state. There appears to be little effect until the final 150 seconds where NO_x is significantly higher in *builds 2 and 3*.

The points noted above have been observed by considering only 4 different setups, however results from all test setups will now be considered. It was seen that during the urban phase of the NEDC, lower NO_x occurred when the engine was hotter because of the engine control strategy acting to retard injection. For each test, total NO_x emissions over phase 1 of the NEDC have been plotted against cumulative engine temperature and cumulative injection timing[63] (figure 5.15). The graphs show good correlation showing a strong relationship between NO_x, the thermal state and injection timing, however because the excitation in injection timing is caused by the changes in warm-up rate, it is not possible to separate the two effects.



(a) *(b)*

Figure 5.15: Urban cycle NO_x correlations with (a) engine temperature and (b) injection timing

In contrast to results over the first stage, over the last 200 seconds (100kph and 120kph cruises), no strong correlations were seen based on engine temperature, injection timing or EGR alone (R^2 values less than 0.3). To understand these effects a linear response model was generated based on these three inputs and a surface generated using the model is shown in figure 5.16. The wireframe surface shows the effect on NO_x of variations in injection timing and temperature at the mean EGR condition. Although neither injection timing nor EGR were directly perturbed, the effect of engine temperature on engine strategy caused variations. A

[63] Whilst these cumulative values have little physical meaning, they do give an estimate of average temperature and timing over the period as a whole. In the case of cumulative temperature, a higher value signified that the engine has operated hotter on average. In the case of cumulative injection timing, a higher number indicates that the engine has run with more advanced timing. Mean values of these variables over the test would show the same relationships, clearly, but the cumulative plots allow a good impression of time varying behavior over the test to be observed and these cumulative values have been used in place of mean values throughout for commonality.

response model was necessary due to the cluster of tests for which injection is slightly retarded (cumulative injection timing -170°BTDC.s) which shows that injection timing has an impact on NO_x emissions in this region. The model also showed that engine temperature has a strong effect on NO_x emissions under these conditions. The effect of EGR has not been shown because there was very little variation.

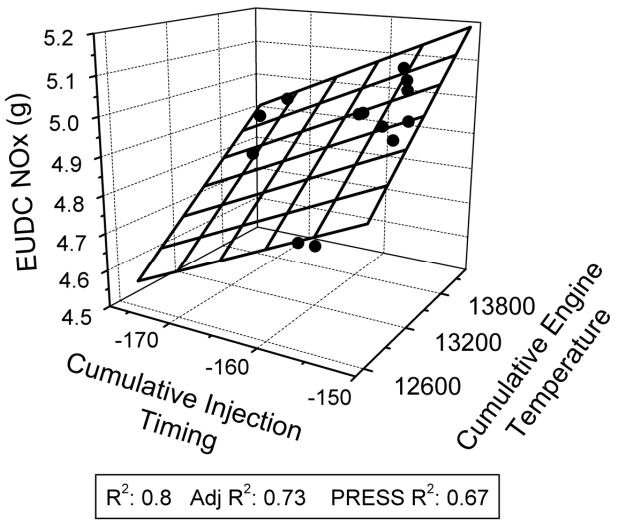


Figure 5.16: Response model for NO_x over 100kph and 120kph cruises of NEDC with respect to injection timing and engine head temperature at mean EGR condition.

4. *Inclusion of injection timing in warm-up strategy*

The calibrations with faster warm-up resulted in benefits both in fuel consumption and NO_x as a result of hotter engine and retarded injection timing. This reduction in NO_x through retarding injection timing can be viewed as compromising the overall potential for reduced fuel consumption and a re-calibration of injection timing could increase the fuel consumption benefits. To assess the potential improvement in fuel consumption under iso-NO_x conditions, a series of tests were conducted in the *Build 3* configuration with varying injection timing over the 4 urban cycles. Three separate tests were conducted during which injection timing was advanced by 5°, 3° and 1°CA, but timing was restored to the default value for the EUDC to limit the impact on NO_x emissions in this phase. The following analysis will therefore concentrate on phase 1 of the NEDC, corresponding roughly to the warm-up period of the engine. Figure 5.17 shows the reference injection timing for the *build 3* calibration and the offset of injection timing of the three other tests. A constant offset is maintained despite the transient nature of the test.

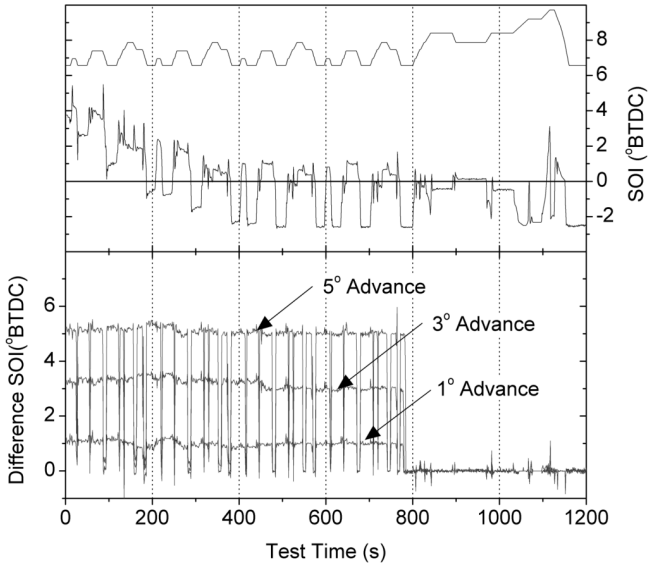
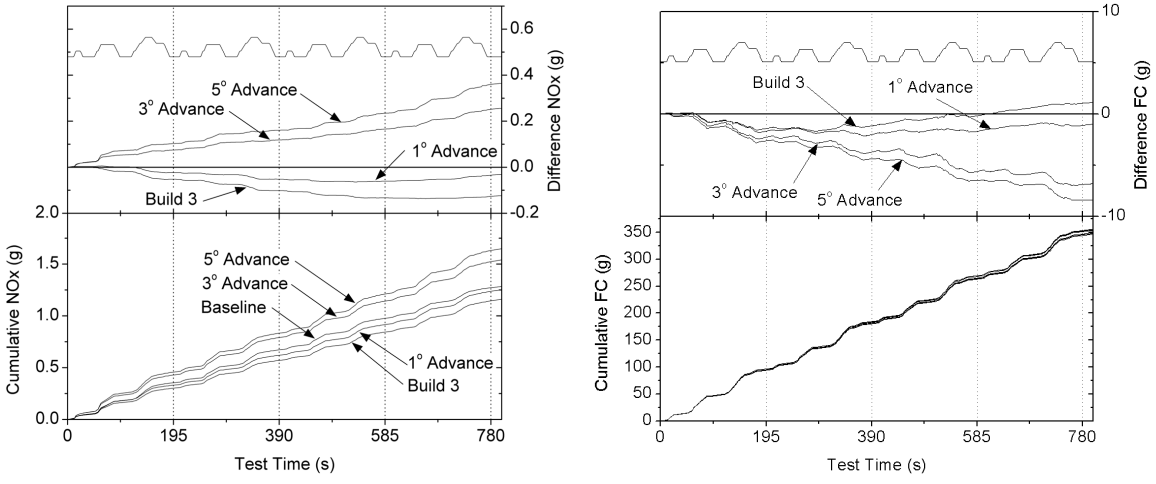


Figure 5.17: Start of injection (SOI) for baseline test and modified injection timings over urban section of drive cycle

The impact of offset injection timing on NO_x emissions and fuel consumption is shown in figure 5.18. These are compared to the *baseline* calibration: as was observed previously in figure 5.12, the effect of faster warm-up results in lower NO_x because of retarded injection. Advancing the timing at these conditions raises the NO_x emissions closer to the *baseline* for 1° advanced, and ultimately increases NO_x for 3° and 5° advance. The opposite effect is seen for fuel consumption.



(a) *(b)*

Figure 5.18: Cumulative (a) NO_x emissions and (b) Fuel consumption over 4 urban cycles and difference to baseline test for each injection timing setting of Build 3

The results over the 4 urban cycles are presented in the form of a NO_x/FC trade-off in figure 5.19. This gives estimate of the potential fuel consumption benefit at iso-NO_x and highlights the need to include the engine strategy in any modifications to the engine cooling circuit. At this stage of the project, the actual benefit in fuel consumption is not of primary concern. The system behaviour demonstrated in this chapter will be used in chapter 7 to organise a more rigorous testing approach to optimise the use of the prototype TMS.

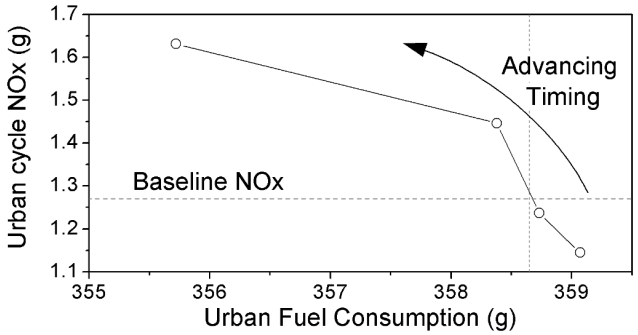


Figure 5.19: NO_x/Fuel consumption trade-off over urban cycles for varying injection timing under build 3 conditions

The shape of the trade-off curve is somewhat unexpected as it is concave with respect to the origin. It would be expected that the shape be convex, as was seen in the results presented by Brace *et al.* [23]. However, the curve shown in figure 5.19 represents only a small section of the overall trade-off curve that would be achieved if timing were varied over a wider range. In this small section of the curve it is reasonable to assume that a linear fit to the data would be a good approximation of the relationship. In fact, when fitting a curve to the data, only a linear fit has any validity, with any higher order fits subject to over-fitting, suggesting that the concave shape of the curve is a function of experimental scatter in the data.

5. *Chapter summary and conclusions*

A number of experimental investigations have been presented in this chapter both under steady state and transient conditions. Engine friction has been measured over NEDC using the indicator method. Steady state experiments have highlighted the engine response to changes in thermal state. Finally the candidate hardware has been used to improve engine warm-up The findings from this chapter may be summarised by the following points:

1) Total engine friction was strongly linked to temperature and engine speed and to a lesser extent to engine load. The relationship was consistent with the on-engine friction regimes reviewed in chapter 2 and other published results.

2) The candidate hardware improved warm-up giving up to 10^{o}C hotter coolant at the end of phase 1. This improvement was primarily due to throttling of engine-out coolant flow which isolated parts of the circuit, effectively reducing thermal inertia during warm up.

3) The faster warm-up rate described in (2) resulted in 2% benefit in fuel consumption and a 3% increase in NO_x. More aggressive throttling gave faster warm-up rates but unexpectedly resulted in an increase in FC and a reduction in NO_x.

4) Steady state experiments highlighted the interactions between the engine control strategy and thermal state, notably with respect to injection timing. Through the same process, injection timing was retarded sooner with enhanced warm-up rates, resulting in lower NO_x emissions. This compromised fuel consumption benefit and explained the unexpected results in (3).

5) The behaviour of the engine during these scoping experiments will be used to establish a suitable experimental design to optimise the use of this prototype system. Notably it is important to include advancing injection timing in any optimisation procedure. This will be covered in chapter 7.

Chapter 6 - *Assessing the fundamental behaviour of the lubrication system*

In parallel to the work presented in chapter 5, in this chapter the behaviour of the lubrication circuit will be shown. A series of scoping experiments were performed and these will form the basis for discussion. The chapter is split into two parts with the first looking at the thermal behaviour of the system and the second part covers effects of varying oil flow rates.

Oil temperatures around the external and internal circuits will be presented. The impact of changes to the external circuit on supply temperature will be investigated. Subsequently an analysis of local internal circuit temperatures will attempt to highlight heat generation from friction and heat loss as the engine warms up.

Both the fixed displacement production oil pump and the variable flow device will be used to assess the impact of varying lubricant flow. In a first stage, the impact on NEDC fuel consumption and emissions will be presented. In a second part, changes in the thermal behaviour of the system will be presented following the impact on piston cooling. Finally indicated results will be used to measure changes in oil pump work.

1. *Thermal behaviour of lubricating circuit*

The thermal behaviour of the engine oil system will be presented in both the internal and external circuits. Both the production and prototype external circuit will be considered, however effects relating to the variable flow oil pump will be discussed in a subsequent section.

Figure 6.1 shows the evolution of oil temperatures for the production engine. The temperature in the sump, at pump discharge (pre filter) and return to engine ladder frame (post filter), are plotted The external circuit of the production engine consists of the oil filter and the oil/coolant heat exchanger; however in this configuration coolant flow is inhibited through this component. The results show that over the initial phase inertia effects of the external circuit cause the post filter temperature to remain low despite an immediate rise in pump delivery temperature. As the engine warms up, there is a small drop in oil temperature over the external circuit until coolant is allowed to flow[64] through the oil cooler at the very end of the drive cycle.

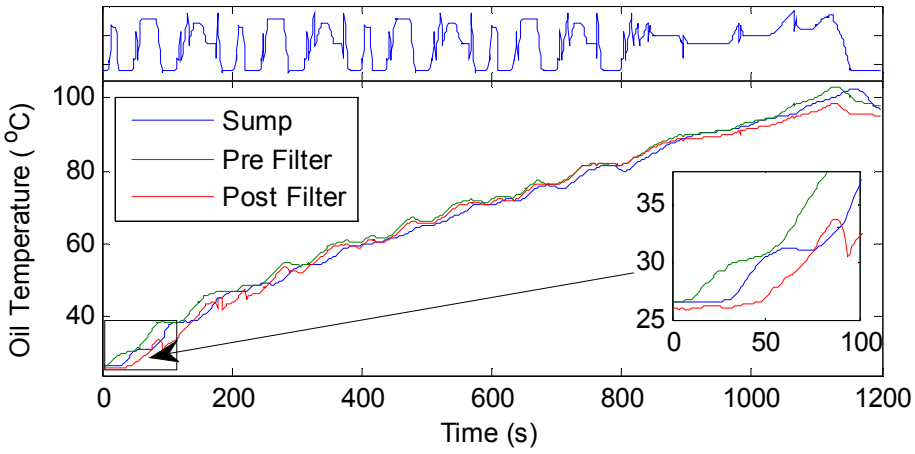


Figure 6.1: External oil circuit temperatures for production engine setup

Figure 6.2 shows sequential oil temperatures in the circuit from the sump to the main bearings over the first 200 seconds. Results are shown for four different configurations of the external circuit using all combinations of oil/coolant heat exchanger and oil-cooled EGR. In all cases the sump and pre-filter temperatures are approximately the same if discrepancies relating to variations in start temperature are ignored. The engine delivery temperature in the main

[64] In the production engine, coolant flow through the oil cooler is controlled by a thermostat (see chapter 4)

gallery rises much faster for both tests using the oil-cooled EGR. The rise is also strongly linked to operating point as the EGR gas temperature will fluctuate depending on engine load. This rise is almost immediate, whereas the other two tests suffer a dead time as shown in figure 6.1. After 200 seconds there are variations of up to 10°C in oil temperature. Despite these large variations in in the main gallery, there is very little variation in main bearing oil film temperatures. This could suggest the additional heat is lost to the engine structure or could be a result of increased friction heat with the colder oil.

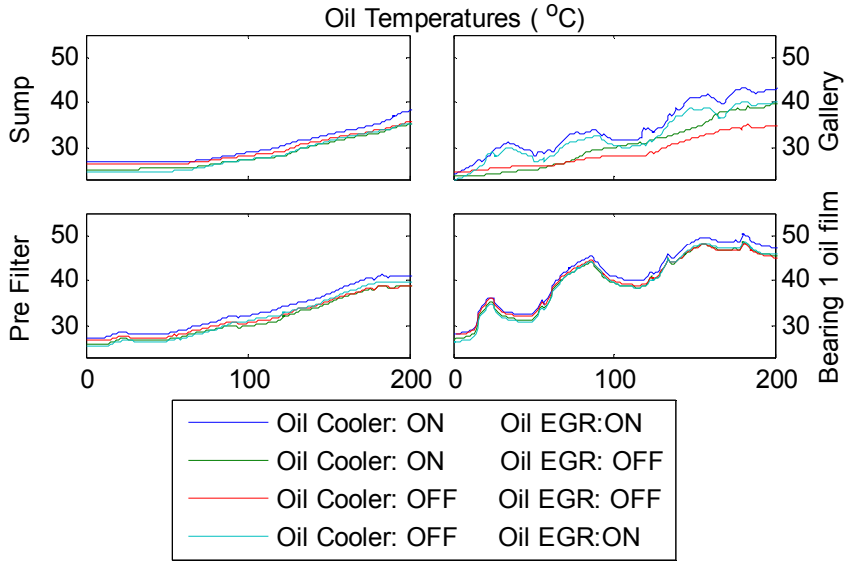


Figure 6.2: Effect of external oil circuit on internal oil circuit temperatures over first 200 seconds of cold-start NEDC

Figure 6.3 shows the same oil temperatures after 500 seconds of the drive cycle. At this point, the test using only the oil/coolant heat exchanger has overtaken the test using only the oil-cooled EGR in terms of main gallery temperature. This is probably explained by the additional energy available in the coolant compared to the EGR gas. Just after start-up, there is very little thermal gradient between the coolant and oil which is unfavourable for heat transfer. As the coolant becomes hotter, a temperature difference develops allowing heat transfer from coolant to oil. In the EGR cooler a favourable thermal gradient is available from the first combustion cycle, however heat transfer is limited by the gas side convection. At this later stage in the drive cycle it should also be noted that the temperature differences in the main gallery *are* reflected in the bearing oil film temperatures.

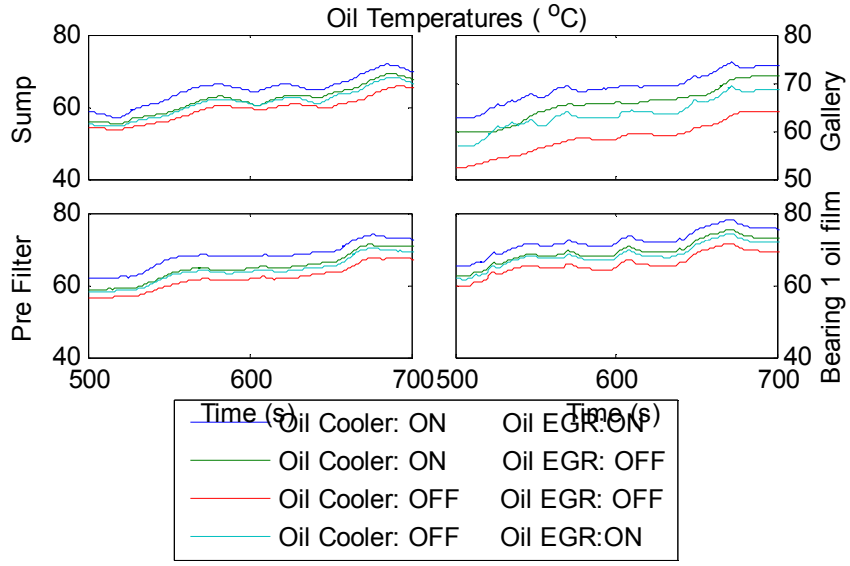


Figure 6.3: Effect of external oil circuit on internal oil circuit temperatures during warm-up of cold-start NEDC

Figure 6.4 shows the oil gallery temperature and the oil film temperatures in each of the main bearings. The oil film temperatures are 5-10°C hotter than the main gallery as a result of the heat generated from friction in the bearings. The difference is significantly higher at the start of the test when oil is colder as engine friction will be higher, generating additional heat. There is also a significant difference for bearing 5 which has a different length to the other bearings.

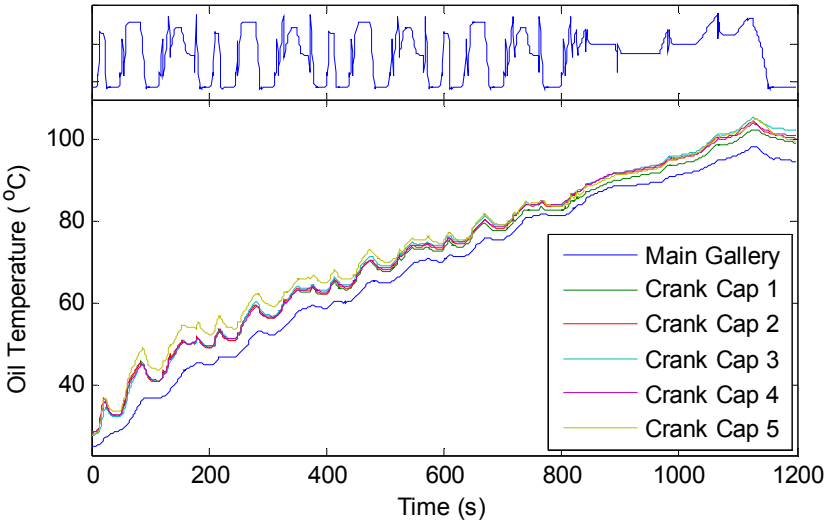


Figure 6.4: Oil gallery and all 5 main bearing oil film temperatures over cold-start NEDC

Figure 6.5 shows oil temperatures measured in the gallery and the oil path in the cylinder head. It is interesting to note during warm-up that there are significant drops in oil temperature of up to 5°C as heat in the oil is transferred to the colder metal structure.

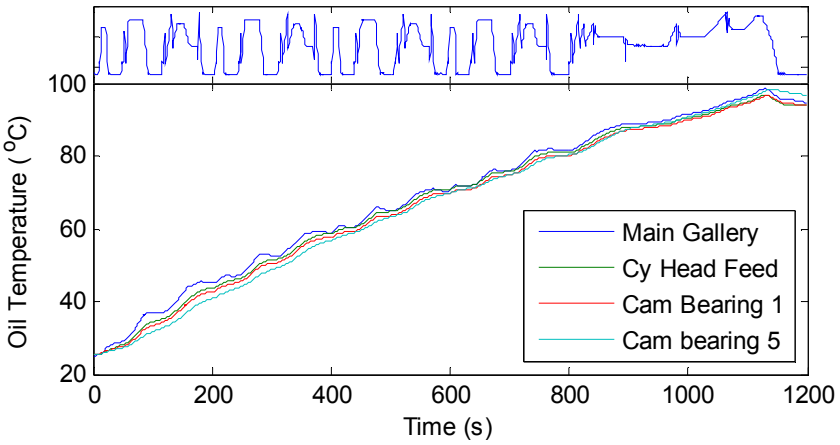


Figure 6.5: Oil temperatures along flow path from main gallery to cam shaft bearings

2. *Engine behaviour under varying oil flows*

2.1. Experimental calibrations

Experiments were performed on the production specification 2.4L engine[65]. Initially the production oil pump was fitted to assess a *baseline* condition, but this was subsequently replaced with the variable flow device. In total, three VFOP calibrations were tested[66] and in each case up to eight valid repeat tests were run to ensure good testing repeatability and confidence in the results presented. When changing from the production oil pump to the VFOP, engine oil was replaced, however both oil batches had previously been aged 10 hours to reduce testing error due to oil aging[67]. The change of VFOP strategy was purely software based and to increase experimental rigour, different configurations were run alternately rather than in batches. The different configurations aimed to control the oil supply pressure in the ladder frame (after the external circuit) to particular set-point; details of the builds are given in table 6.1.

[65] Initially a near production cooling circuit was used with flow meters and the oil cooler thermostat removed.

[66] In subsequent testing, it was required to provide a high oil flow rate without physical removal of the variable flow oil pump to avoid disturbing the system. The variable flow device was run a maximum displacement which approximated the behaviour of the production pump.

[67] Generally fuel consumption of the powertrain will improve over the first few hours following an oil change due to initial reductions in viscosity and possible additive interactions [151].

Build	Oil pump	Target oil pressure
1	Production	Max
2	VFOP	3bar
3	VFOP	1-2bar

Table 6.1: Three build setups for main experimental campaign

The production oil pump does not require any additional control, however the VFOP used Ati Vision software to determine set-point using a 2D map based strategy with engine speed and fuelling as inputs. Post filter pressure feedback control was also used as described in chapter 4. The effects of oil pump type and configuration on *pump delivery pressure* (pre-external circuit) and *engine delivery pressure* (post-external circuit) over the oil temperature and engine speed range are shown in figures 6.6 and 6.7 respectively[68]. The production oil pump delivers oil pressures over 6 bar when cold which gradually drop to around 4bar when the oil has warmed up. The VFOP demonstrates good control of oil pressures both with a constant 3bar target and an optimised 1-2bar calibration. As the lubricant warms up the pump is unable to meet the 3bar demand under idling conditions and supplied a pressure similar to that of the production pump.

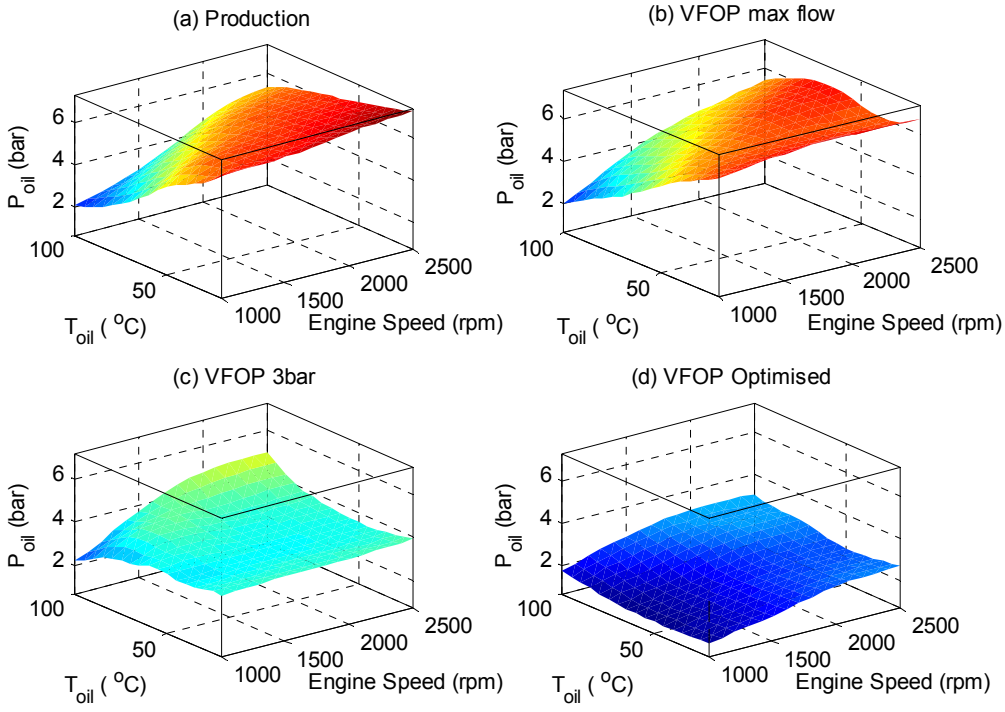


Figure 6.6: Oil pump delivery pressure (pre external circuit) with oil temperature and engine speed for (a) production oil pump, (b) VFOP at max flow, (c) VFOP with target pressure 3bar and (d) VFOP optimised

[68] The surfaces in both figures have been obtained from response models. RBF network models were used because of the complex response nature of the pressure control algorithm [152].

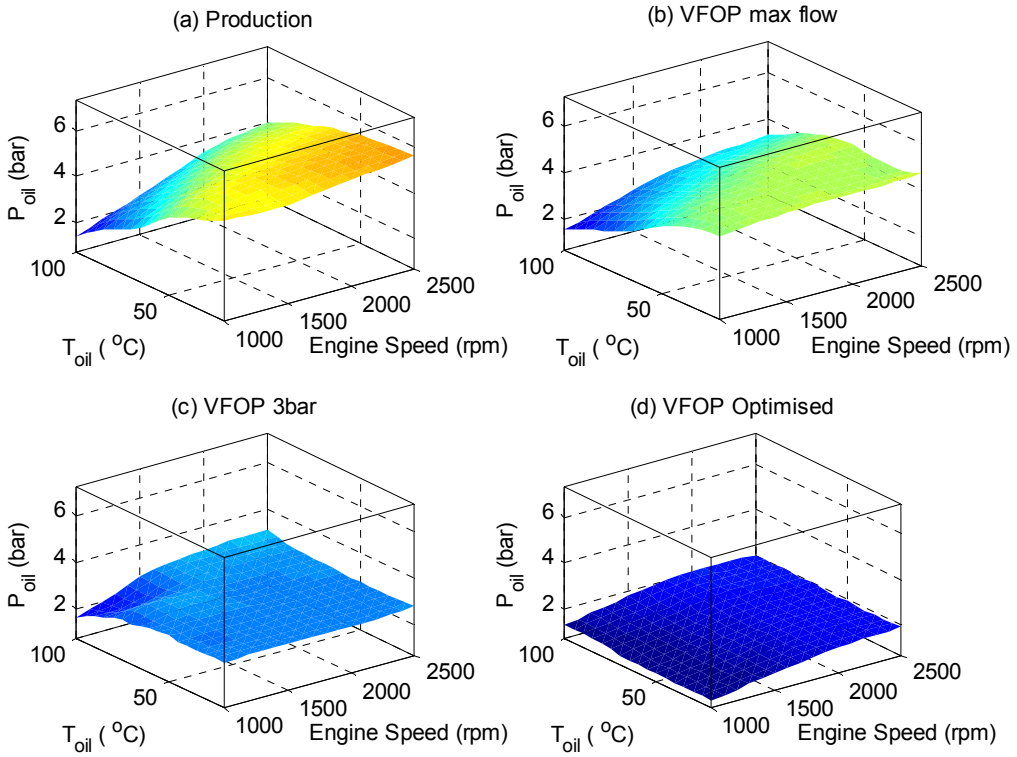


Figure 6.7: Engine oil delivery pressure (post external circuit) with oil temperature and engine speed for (a) production oil pump, (b) VFOP at max flow, (c) VFOP with target pressure 3bar and (d) VFOP optimised

During the initial testing, in-cylinder pressure data was not available and subsequent tests were run when this equipment had been installed. At this stage the layout of the coolant circuit had been modified in-line with other research, and it was not possible to return to the production oil pump. However, the VFOP was varied over its full operating range giving similar behaviour to the production pump (see figure 6.6).

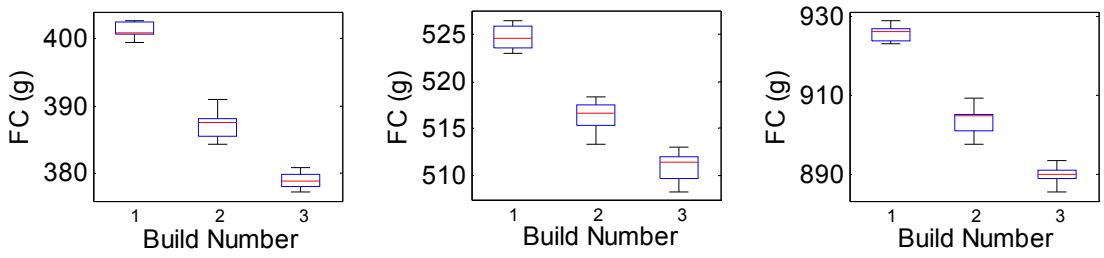
2.2. Results

2.2.1. Macroscopic results

(a) Fuel consumption

Fuel consumption measurements for all tests are presented in the form of box plots[69] for each phase of the NEDC cycle and for cold- and hot-start tests in figures 6.8 and 6.9 respectively (build numbers refer to those described in table 6.1). These results show good repeatability with a typical coefficient of variation of around 0.5%. The fuel consumption results presented are from the carbon balance method, however these agreed well with gravimetric fuel balance measurements when the correction procedures described in chapter 3 are applied.

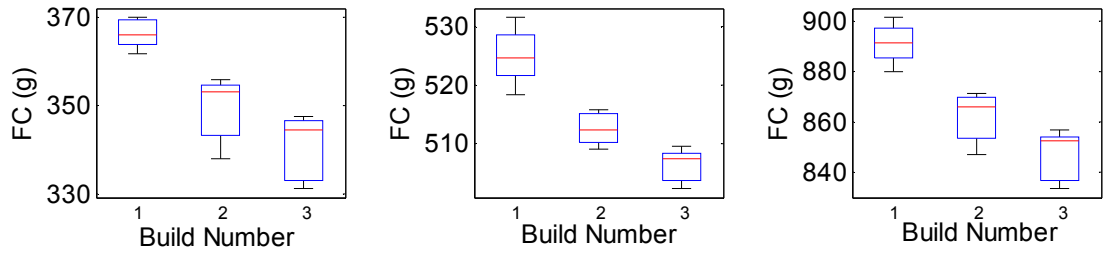
Significant reductions in fuel consumption were observed in both the cold-start and hot-start tests by replacing the oil pump and lowering the oil pressure in builds 2 and 3. The lowest oil pressure set-point in *build 3* offered a 36g (4%) benefit over the NEDC for cold tests. This can be split into 22g (6%) during phase 1 and 14g (3%) over phase 2. Similar results are observed from the hot-start tests which suggests the discrepancies between phase 1 and phase 2 are not predominantly due to changes in oil viscosity. Close observation of the speed/torque trace for the NEDC cycle shows that despite phase 1 representing the lower power phase, engine speeds tend to be higher over this period, meaning larger oil pumping work. These results agree with those published by Brace et al. [79] which used the same pump on a different engine.



(a) Phase 1 *(b) Phase 2* *(c) NEDC*

Figure 6.8: Cold-start fuel consumption for (a) phase 1, (b) phase 2 and (c) NEDC

[69] The box plots show the median (red line), upper and lower quartiles (box edges), minimum and maximum (whiskers) and show any outliers where appropriate. Outliers were detected as being 1.5 times the inter-quartile range from to edges of the box.



(a) Phase 1 *(b) Phase 2* *(c) NEDC*

Figure 6.9: Hot-start fuel consumption for (a) phase 1, (b) phase 2 and (c) NEDC

(b) Feed-gas emissions

NO_x, HC and CO emissions, sampled between the turbocharger and the catalyst, for phases 1 and 2 are shown for cold- and hot-start tests in figures 6.10 and 6.11 respectively. The spread of data is larger than fuel consumption showing less repeatability for these measurements, however trends are still apparent. In most conditions NO_x emissions are up to 3% higher with reduced oil flow. The exception to this is phase 1 of the cold drive cycle where NO_x emissions are 5% lower. HC and CO emissions follow similar trends throughout and tend to reduce by 3% to 5% with lower oil flow. Phase 1 of the cold-start drive cycle shows the smallest effect whilst during the other phases when the engine is running hot, larger reductions are seen.

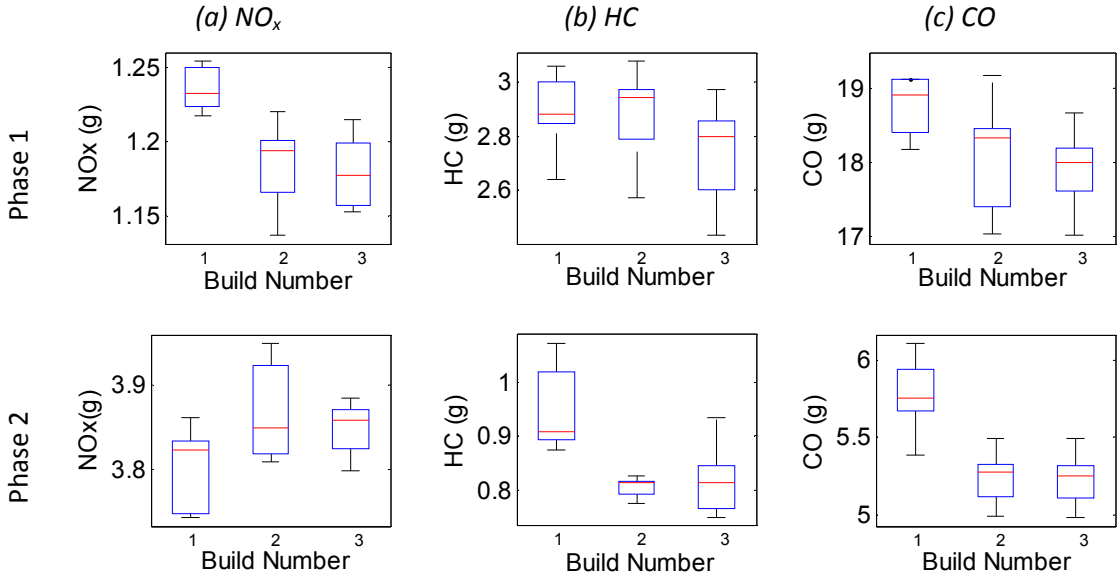


Figure 6.10: Cold-start NO_x (a), HC (b) and CO (c) feedgas emissions for phase 1 and 2

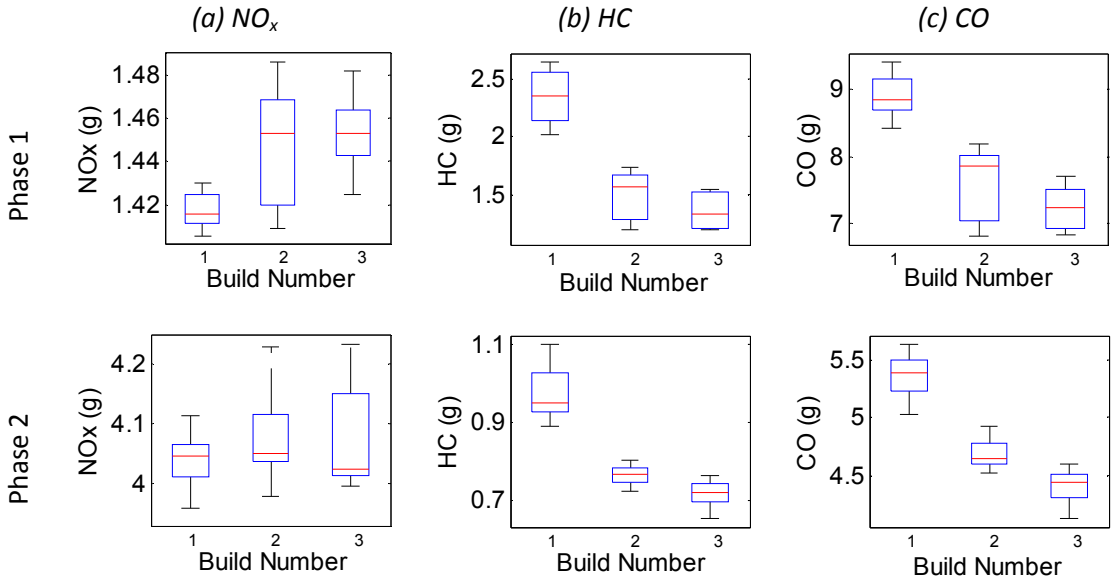


Figure 6.11: Hot-start NO_x (a), HC (b) and CO (c) feedgas emissions for phase 1 and 2

2.2.2. Engine thermal behaviour

Results in this section will cover the phenomenological effects observed in the engine. Of key importance for the emissions is the thermal state of the combustion chamber. Such is the design of the engine, piston cooling jets use engine oil to cool the piston crown and at the same time the combustion chamber walls. Instrumentation of the piston itself is beyond the scope of this thesis, an insight into these effects will be given by assessing temperatures around the engine.

(a) Oil and bearing Temperatures

Figure 6.12 shows main gallery oil temperature for different oil pumping pressures and for both hot- and cold-start tests. It is clear from the cold-start tests that a reduced oil flow leads to reduced oil temperatures in the main gallery by up to 4°C during warm-up. This may be explained by reduced work input to the lubricant and less heat generation in the pressure relief valve, but also because of less overall heat transfer to the oil around the engine, notable in the piston and bore regions because of reduced flow from piston cooling jets. The reduced flow would also affect the heat transfer in the oil cooler, and a lower flow would be less favourable. However, inspection of energy loss from coolant in the oil cooler shows that there is increased heat transfer with lower oil flow. This is explained by the larger temperature gradient between coolant and the colder oil in these conditions. Results from the hot test would appear to confirm this as well because lower oil flows again yield coldest lubricant temperatures. If heat transfer was severely affected by oil flow, it would be expected that the

lubricant be hotter in hot-start tests due to less effective cooling. However, there is little difference in oil gallery temperature during the hot tests, with the lower flow approximately 1°C cooler than the production pump.

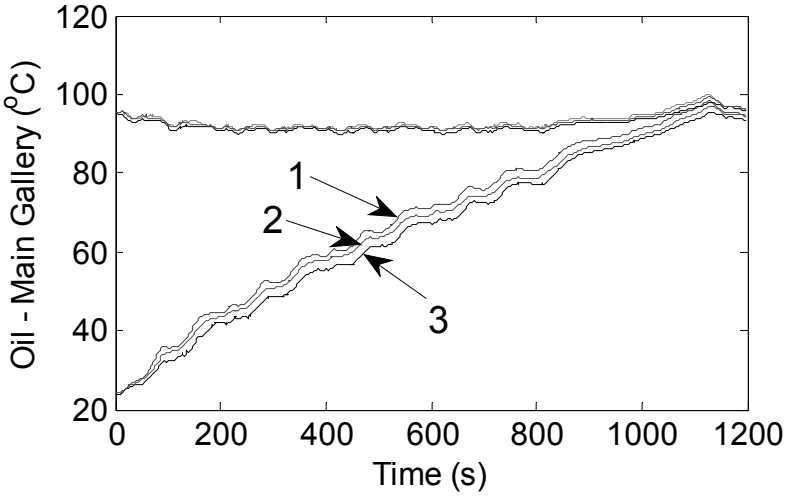


Figure 6.12: Main Gallery oil temperature for builds 1, 2 and 3 for hot and cold-start tests

The knock on effect in the cold-start tests is that crankshaft and cam shaft bearing oil film temperatures were lower in the case of lower oil flow. In the host start tests, the oil temperature at the cam shaft bearing film correlates well with that in the main gallery, with the higher oil flow giving higher temperatures (figure 6.13 (a)). However, the converse is true of the main crank bearings: the lower oil flow, with lower main gallery oil temperatures, gives higher oil film temperature (figure 6.13 (b)).However these differences remain small in the hot-start NEDC.

(b) Cylinder liner temperatures

Many temperatures were measured around the cylinder liners of cylinders 2 and 3 (see chapter 4) giving a comprehensive description of the thermal state of the upper section of the engine block. A selection of results[70] showing the key differences over the first 600 seconds of the cold-start NEDC are shown in figure 6.14. These include temperatures down the liner, at different depths and on intake and exhaust sides. A small diagram is also shown indicating the locations of the thermocouples.

[70] These temperatures have been selected as they show the dominating trends.

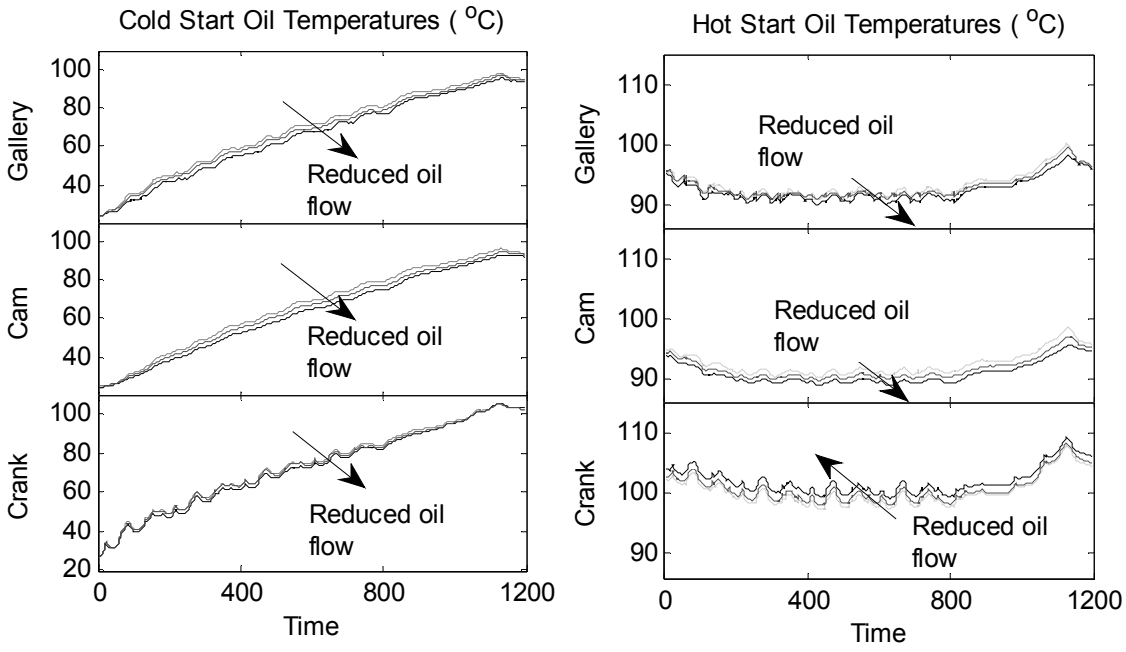
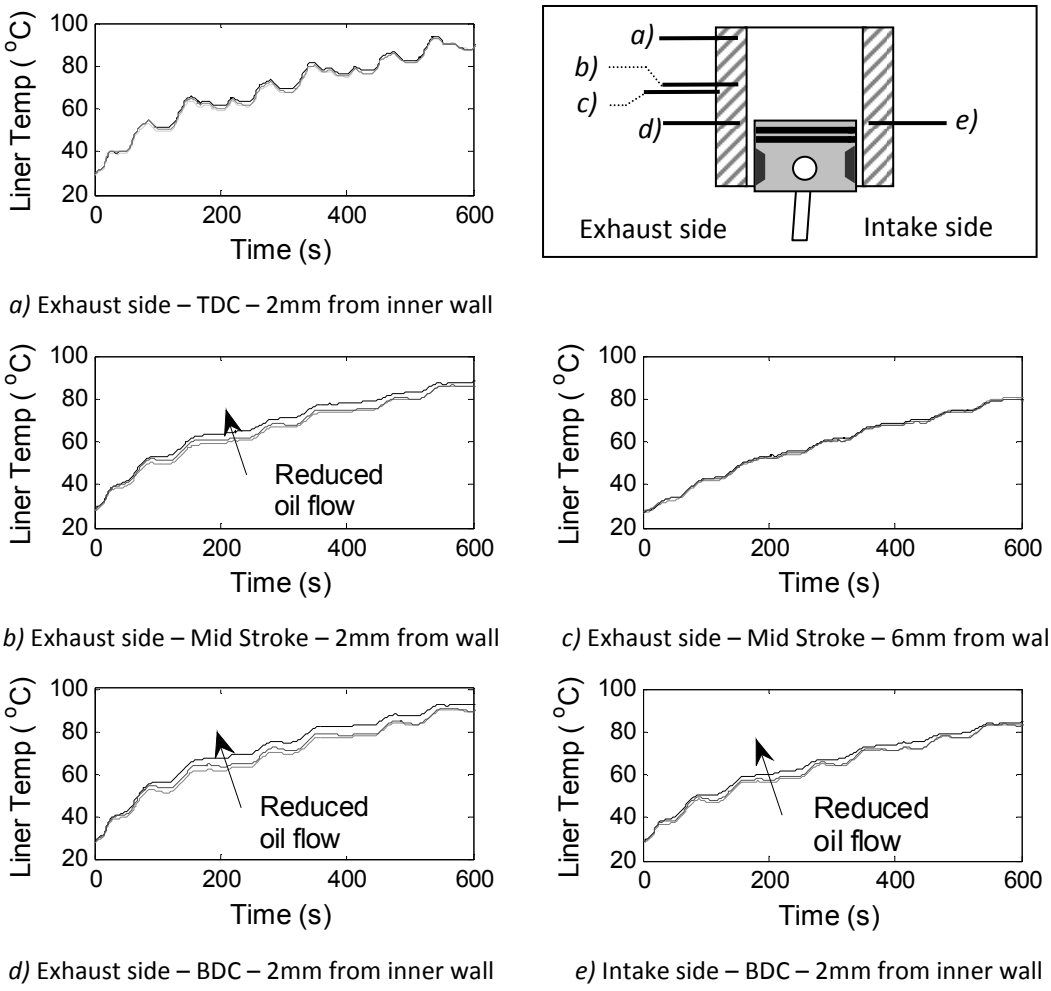


Figure 6.13: Oil film temperatures for (a) cold-start and (b) hot-start tests

From the temperature profiles in figure 6.14 it is clear that the reduction of oil flow does have an effect on cylinder liner temperatures. These are most marked in the lower half of the stroke, closer to the combustion chamber (figure 6.14 b, d and e). After 200 seconds, lower liner temperatures in the test with lowest oil flow lead the production setup by 4-6°C. Temperatures at the top of the cylinder are much less affected by the oil flow (figure 6.14 a); in this position after 200 seconds the lowest oil flow setup leads the production oil pump by only 2°C. Deeper into the liner, away from the combustion chamber temperatures are also less affected by oil flow rate (figure 6.14 c) with a difference of less than 1°C between setups. Temperatures nearer bottom of the liner and further away from the coolant jacket would be expected to be more dependent on oil cooling effects. The variations in liner temperatures here are of similar order or magnitude to the simulated piston temperature changes reported by Agarwal and Varghese [80, 81].

In addition to differences down the cylinder bore, there are also differences comparing exhaust and intake sides of the engine (figure 6.14 d and e). On the intake side, after 200 seconds the low flow test leads the production build by 4°C, whereas on the exhaust side at this same time the difference is 6°C. In all tests, after 600 seconds the temperatures on intake side appear to be converging as the engine warms up, however on the exhaust side the higher temperature prevails longer. These differences may be a result the different piston forces acting on the liner on the thrust and anti-thrust sides.



a) Exhaust side – TDC – 2mm from inner wall

b) Exhaust side – Mid Stroke – 2mm from wall

c) Exhaust side – Mid Stroke – 6mm from wall

d) Exhaust side – BDC – 2mm from inner wall

e) Intake side – BDC – 2mm from inner wall

Figure 6.14: Selected liner temperatures over first 600 seconds of NEDC for cold-start tests. A to E represent temperatures at locations detailed in cylinder diagram

Figure 6.15 shows the same temperature measurement as figure 6.14 d), but for both hot and cold-start tests and for the whole NEDC. Despite a convergence of liner temperature at around 800 seconds in the cold-start tests, under fully warm conditions, the experiments with lower oil flow operate 2-6°C hotter than with the production oil pump. This convergence of temperatures for each test is curious and further explanation may be found by looking at coolant temperatures.

Despite the significant effects on oil and liner temperatures presented above, the oil flow has little effect on coolant warm-up rate over the first 400 seconds of the drive cycle (see figure 6.16). After this, and until the coolant reaches fully warm operating temperature, the setups with lower oil flow cause slightly lower coolant temperatures (up to 3°C). Analysis of coolant temperatures around the circuit shows that this is primarily due to a lower temperature of coolant exiting the oil cooler as a result of lower oil temperatures (see figure 6.12). This

variation in coolant temperatures occurs simultaneously with the convergence of liner temperatures during warm-up. This suggests that it is the higher coolant temperature that causes the slight increase in liner temperatures for the production setup and not changes in piston cooling. When the coolant reaches approximately 90°C the thermostat opens and regulates to a constant temperature, irrespective of oil flow. At this point, differences in liner temperatures due to oil flow are once again apparent. Local coolant temperatures within the cooling jacket confirm this with the largest effect on the air intake side. This is closer to the coolant inlet[71] and would partially explain the liner temperature differences observed in figure 6.14 d and e.

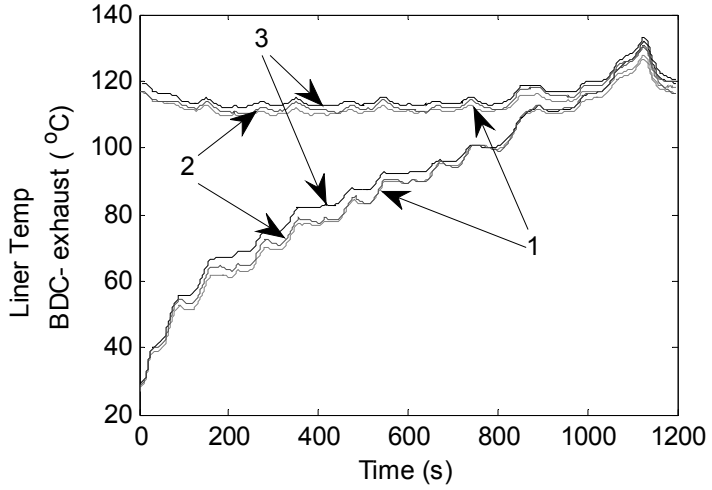


Figure 6.15: Cylinder liner temperature for NEDC at BDC – exhaust side – 2mm from inner wall

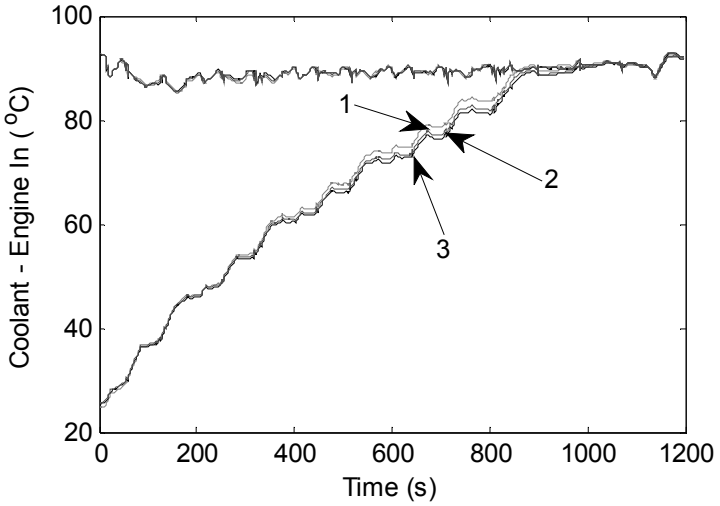


Figure 6.16: Coolant inlet temperature for hot and cold-start NEDC for builds 1, 2 and 3

[71] The cooling system is a cross-flow system with coolant entering the block from the intake side via a coolant rail to distribute the flow evenly over each cylinder.

2.2.3. Indicated results and pumping work

Further tests were completed after the initial investigation to gather in-cylinder pressure data using the VFOP in three calibration conditions: max flow, 3bar target pressure and 1-2bar target pressure. Figure 6.17 shows friction and accessory power and cumulative work over the NEDC cycle. The power plot is difficult to assess due to significant noise present during the transient phases[72]. Integrating this signal to estimate work allows easier assessment of the results. The differences resulting from variations in oil flow are small in the context of overall friction and accessory loading. However there is a trend of reduced work with reduced oil flow.

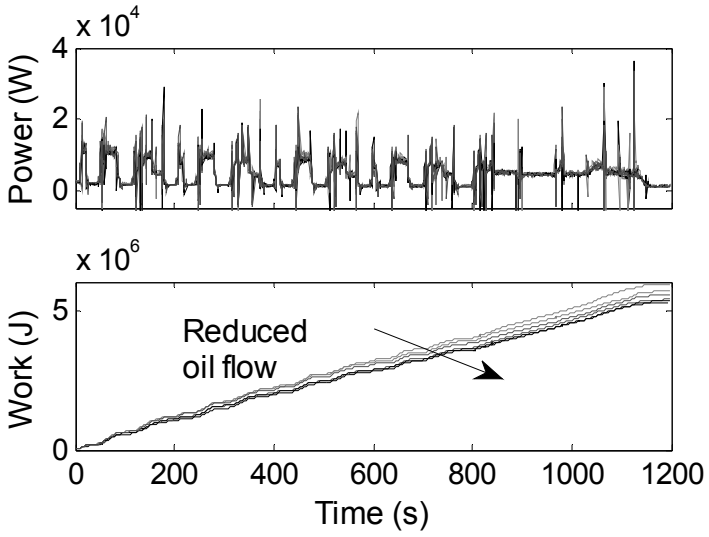


Figure 6.17: Friction and accessory power and work from NEDC assessed from difference between indicated and brake mean effective pressures

Figure 6.18 shows a plot of total friction and accessory work over the NEDC against the mean post-filter oil pressure. There is a clear correlation of data with the higher oil pressures leading to higher overall work. However, whilst the ranking order is consistent, there are significant test to test variations. In depth analysis of the test data shows that there is no difference in warm-up rates to explain the scatter in results. However, there are differences in battery charge over the cycle that correlate with the variations and may explain the measurement scatter.

[72] This is partly a result of time alignment of the two signals from different acquisition systems but also due to the different operating frequencies of these two systems.

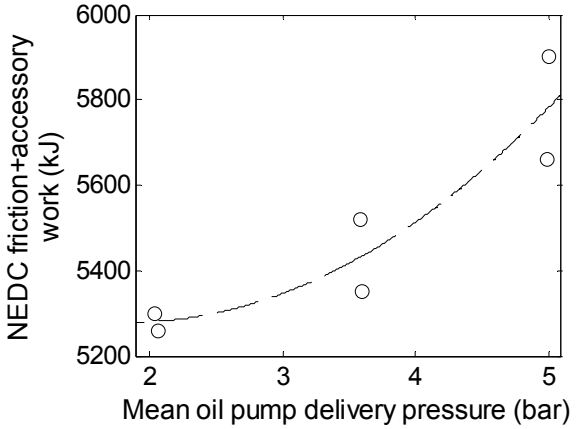


Figure 6.18: NEDC total friction and accessory work vs. mean oil pump delivery pressure

2.3. Discussion

The introduction of a variable flow oil pump offers an additional control parameter for the engine. Initial results demonstrated a clear benefit in fuel consumption from the lower oil flow rates. Subsequent in-depth analysis of the engine operating behaviour has demonstrated that there are some direct effects of reducing the oil flow which include lower oil pump work. Some secondary effects on oil, coolant and metal temperatures were also measured. Each of these will have different effects on fuel consumption and emissions.

The reduced oil pumping work will reduce fuel consumption. Figure 6.19 shows the changes in indicated work against changes in fuel energy consumption over the NEDC. There is a clear correlation between the two showing the strong effect of reduced accessory load or friction. The majority of these points lie close to the *y=x* line showing that reduced pumping work accounts for the majority of fuel consumption benefits. The offsets from this line will be accounted for by impacts on the thermal state of the lubrication circuit and the combustion process. The lower temperatures in the bearings during warm-up will tend to increase friction whilst the higher liner temperatures will tend to reduce it. The change in oil flow may also impact on the lubrication regimes in the bearings with more or less lubricant present. It is difficult in the context of a firing production engine to explain further any of these effects and further measurements on a dedicated experimental facility are recommended.

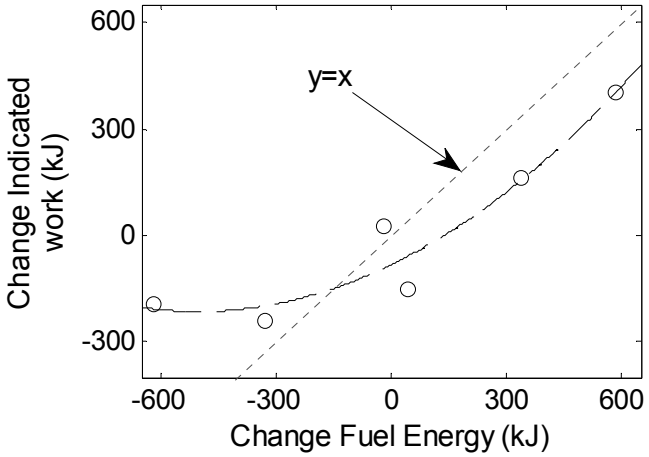


Figure 6.19: Changes in fuel consumption energy against indicated energy over NEDC for different oil pump flow rates (0kJ represents mean test results)

Liner temperatures were highest in all tests with lowest oil flow. Under hot conditions, these tests also produced the highest NO_x emissions and lowest HC and CO emissions (Hot tests and phase 2 cold tests). When the engine is operating fully warm there are no discrepancies in the injection timing meaning the differences in emissions can be largely attributed to the differences in temperatures. However, during warm-up NO_x are reduced with low oil flow and there is little effect on HC and CO emissions. This may be a result of the lower fuel consumption that reduces combustion temperatures.

The effect of adding the dual EGR system increased the pressure drop over the external circuit as shown in figure 6.20. Further system analysis in the context of a production system may require a larger displacement pump under the most extreme operating conditions. If this is the case, then the benefits from a variable displacement device over its fixed displacement counterpart would be even greater.

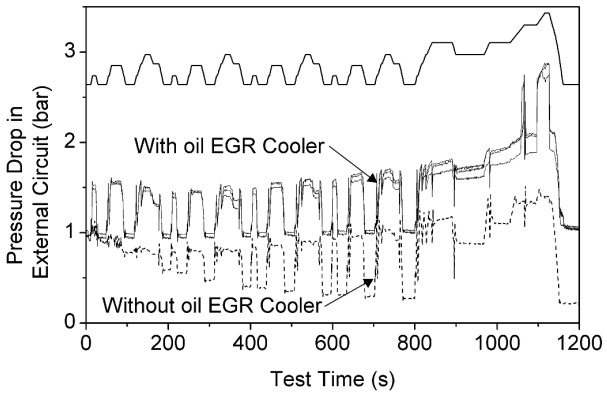


Figure 6.20: Oil pressure drop in external circuit with and without oil EGR cooler fitted

3. *Chapter summary and conclusions*

A number of experiments investigating changes in the lubricant circuit have been performed. Using the specialised instrumentation described in chapter 4, local oil temperatures have given information relating to the thermal behaviour to complement effects on fuel consumption and emissions. Along with the results from chapter 5, these results will be considered when designing the experiment in chapter 7 for system optimisation. From the results and discussion presented in this chapter, the following conclusions may be drawn:

1) In the production engine the external oil circuit comprising of the oil filter and oil cooler (without coolant flow) presents a significant thermal inertia. This inertia effect can be overcome using oil-cooled EGR because the exhaust gases present a high temperature heat source that is immediately available. Later in the cycle, the use of oil/coolant heat exchanger offers better oil warm-up because of a higher energy supply.

2) Local increases in oil temperature of about 10^{o}C were measured in the main bearings compared to the oil gallery as a result of friction heat. In contrast, a drop in oil temperature of 5^{o}C was measured from the main gallery to the valve train supply. Heat addition in the external circuit is not reflected in main bearing oil film temperatures over the first 200 seconds, but is apparent later in the cycle.

3) Engine oil flow also affects the oil supply temperature because of reduced pumping work. Temperatures at the lowest allowable oil flow were 4^{o}C colder than with the production pump over the cold-start NEDC.

4) The thermal effects of oil flow variation affect emissions formation because of changes to piston cooling. Under fully warm conditions with lowest oil flow, NO_x emissions were 3% higher but HC and CO reduced by 3-5% compared to a production setup.

5) The reduction in pump driving torque has a greater effect on fuel consumption compared to adverse effects on engine warm-up from reduced oil flow. The measured fuel consumption benefit over NEDC was 4% but benefits were dependent on engine operating speed, with larger benefits at higher speeds.

Chapter 7 - *Detailed understanding and optimisation of active TMS*

In this chapter a design of experiments approach will be taken to fully investigate the behaviour and interactions of the prototype hardware. The results will subsequently be used to provide optimised calibrations for minimum fuel consumption over the cold-start NEDC. The system viability under fully warm conditions will also be assessed.

The knowledge acquired over the previous two chapters will be used to establish which parameters should be included in the experimental design. Each test point of the designed experiment will represent a cold-start NEDC so that dynamic behaviour may also be captured. With a detailed understanding of the system behaviour in the form of statistical models, optimisation techniques may be applied to establish best use of the additional hardware. These optimised calibrations will subsequently be validated experimentally.

In the second section of this chapter, the suitability of the system to operate under fully-warm conditions will be assessed. This work will focus on benefits from coolant flow throttling (to increased engine temperature at part load) and the performance of oil-cooled EGR. If minimal penalties or further benefits can be demonstrated, they will enhance the case for adoption of these technologies.

1. *System warm-up optimisation*

1.1. Experimental design

A number of different hardware devices were presented in chapter 4 and integrated into the production engine setup. A basic understanding of the way these devices work and interact with the engine was established in chapters 5 and 6. This chapter will build on this understanding using the instrumented engine and aim to optimise the combined use of the new devices. Based on this previous work, the control variables of interest and their different settings[73] are listed in table 7.1. These include the four degrees of freedom added by the active cooling system analysed in chapter 5. Injection timing is included due to interactions between the ECU and changes in engine warm-up rate however the range is reduced from 5^{o} to 3^{o} advance based on results from chapter 5. The VFOP was not included in this experimental design because of the limited effect on thermal behaviour[74] compared to the significant reduction in fuel consumption demonstrated in chapter 6. For all experiments in this section the VFOP was calibrated to minimum flow.

#	Input	Setting 1	Setting 2	Setting 3
1	Engine-out coolant throttle	Mapped	Mapped + min limit	Open
2	Oil cooler Use	On		Bypassed
3	EGR coolant throttle	Low 20%	Mid: 28%	High: 100%
4	EGR Cooler type	Coolant		Oil
5	SOI	Standard	+1.5^{o}	+3^{o}

Table 7.1: Input variables for engine optimisation

A design of experiments approach was used and 17 different calibrations[75] were tested, based on a D-optimal test design; the calibrations are listed in table 7.2. Each test point represented a cold-start (25^{o}C) NEDC and the run order was randomised. The experiments were completed over a 3 week period with two tests per day: the first following an overnight soak, the latter following a forced cool down procedure. Repeat tests showed no significant trend between overnight and forced cool down tests.

[73] Some of the parameters had simple two-level control, for example on/off, whereas other parameters offered more continuous control and were assigned an intermediate level to detect any curvature in the response.

[74] Although it could have been useful to include the VFOP in the experimental design, the case for keeping the number of tests within a manageable time frame meant this factor was omitted.

[75] The total number of tests was a compromise between exploring the multidimensional design space and avoiding a long running testing campaign with possibilities of mechanical failure and drift.

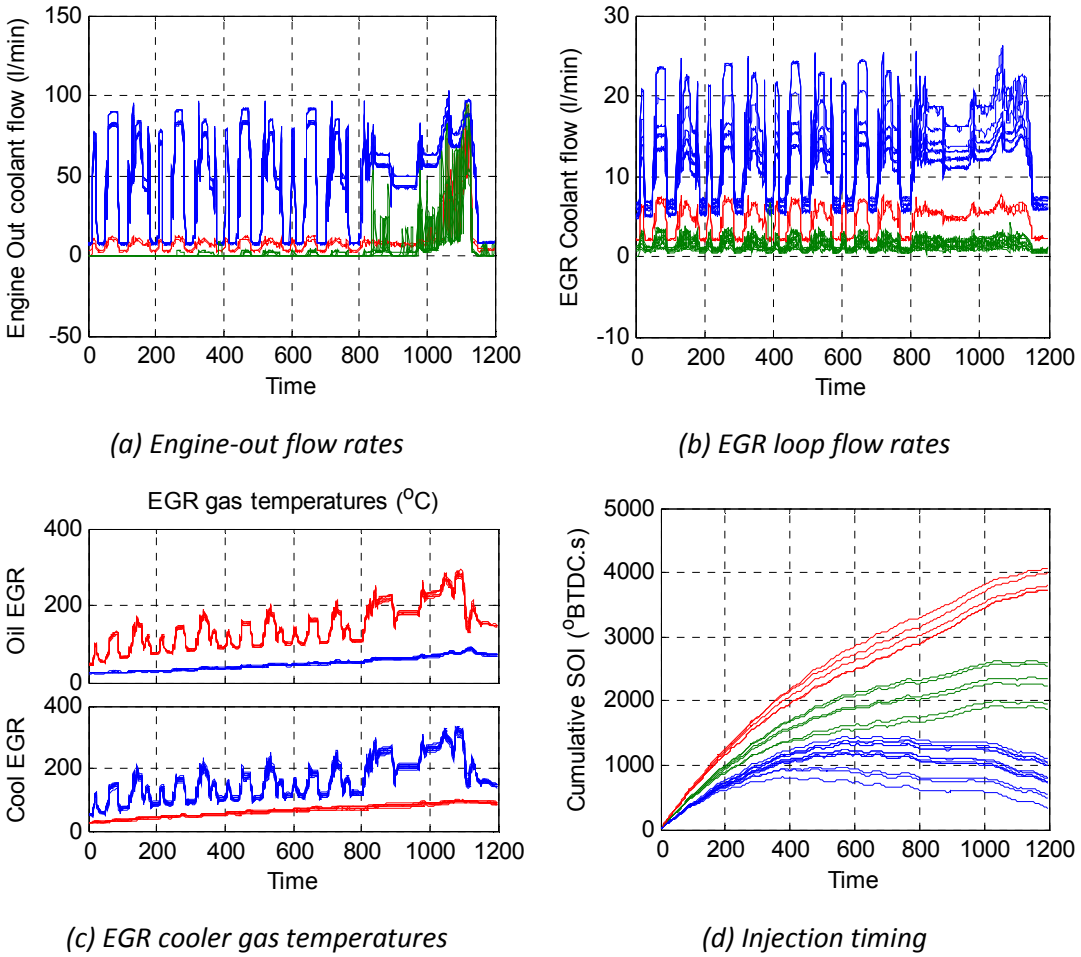
#	1	2	3	4	5
DoE Number	**Eng out coolant throttle**	**EGR Coolant flow**	**Oil Cooler HE Bypass**	**EGR valve**	**SOI advance (°)**
1	Mapped	Low	Bypass	Oil	0
2	Open	Low	Bypass	Oil	1.5
3	Mapped	Low	Bypass	Coolant	0
4	Open	Low	Bypass	Coolant	3
5	Open	Low	On	Oil	0
6	Open	High	Bypass	Oil	1.5
7	Mapped	High	Bypass	Coolant	1.5
8	Open	High	On	Coolant	0
9	Open	Low	Bypass	Coolant	1.5
10	Mapped	Low	On	Coolant	1.5
11	Mapped	High	Bypass	Oil	3
12	Mapped	Low	On	Oil	3
13	Mapped	High	On	Oil	1.5
14	Open	High	Bypass	Oil	0
15	Open	High	On	Coolant	3
16	Mapped	High	Bypass	Coolant	0
17	Mapped*	Mid	On	Coolant	0

* Test 17 used the same mapped setting, but a lower threshold was introduced, effectively maintaining a minimum throttle opening and minimum coolant flow. This produced a result with a mid-flow setting throughout the test.

Table 7.2: Experimental design test points for assessing active thermal management system

Good control of the input variables was demonstrated: Figure 7.1 shows time series plots for all experiments showing the effect of each input variable. The experiments are grouped by colour for the respective setting of each variable (detailed in table 7.1):

- Engine-out (a) and EGR loop (b) coolant flow clearly shows three groups for “low”, “mid” and “high” flows, although there is some scatter due to interactions with other control variables.
- EGR gas temperatures (c) show that only one of the two coolers is working in any one test. It is interesting to note the lower EGR temperatures with oil-cooled EGR compared to coolant-cooled EGR.
- Injection timing, shown in (d) as cumulative injection timing, shows three distinct groups. The variation within these groups is a result of different warm-up rates which affects the progression through the temperature based maps.



(a) Engine-out flow rates

(b) EGR loop flow rates

(c) EGR cooler gas temperatures

(d) Injection timing

Figure 7.1: Variation in control variables for DoE test points for (a) engine out coolant flows, (b) EGR loop coolant flows, (c) EGR cooler type and (d) injection timing

1.2. Results

1.2.1. Raw data assessment

The fundamental behaviour of the system was the same as that presented in chapter 5, showing that the transferral of cooling hardware to the instrumented engine did not have a significant effect on system performance. Because of the large number of input variables compared to the relatively few number of tests, it is not easy to pick out trends from the raw data and the majority of the work in this chapter will rely on statistical response models.

It is interesting to compare the engine behaviour over the DoE experiments with the baseline engine without the additional hardware[76]. Figure 7.2 shows selected oil temperatures for this *baseline* calibration[77] and the spread of measures obtained during the DoE testing. The baseline temperatures lie towards the upper part of the DoE spread over the first 800 seconds showing the main effect of the additional hardware is to reduce warm-up through additional inertia. If the prototype nature of the installation is considered, then benefits should be assessed based on a *baseline* calibration *with* the additional hardware installed but *not* in use.

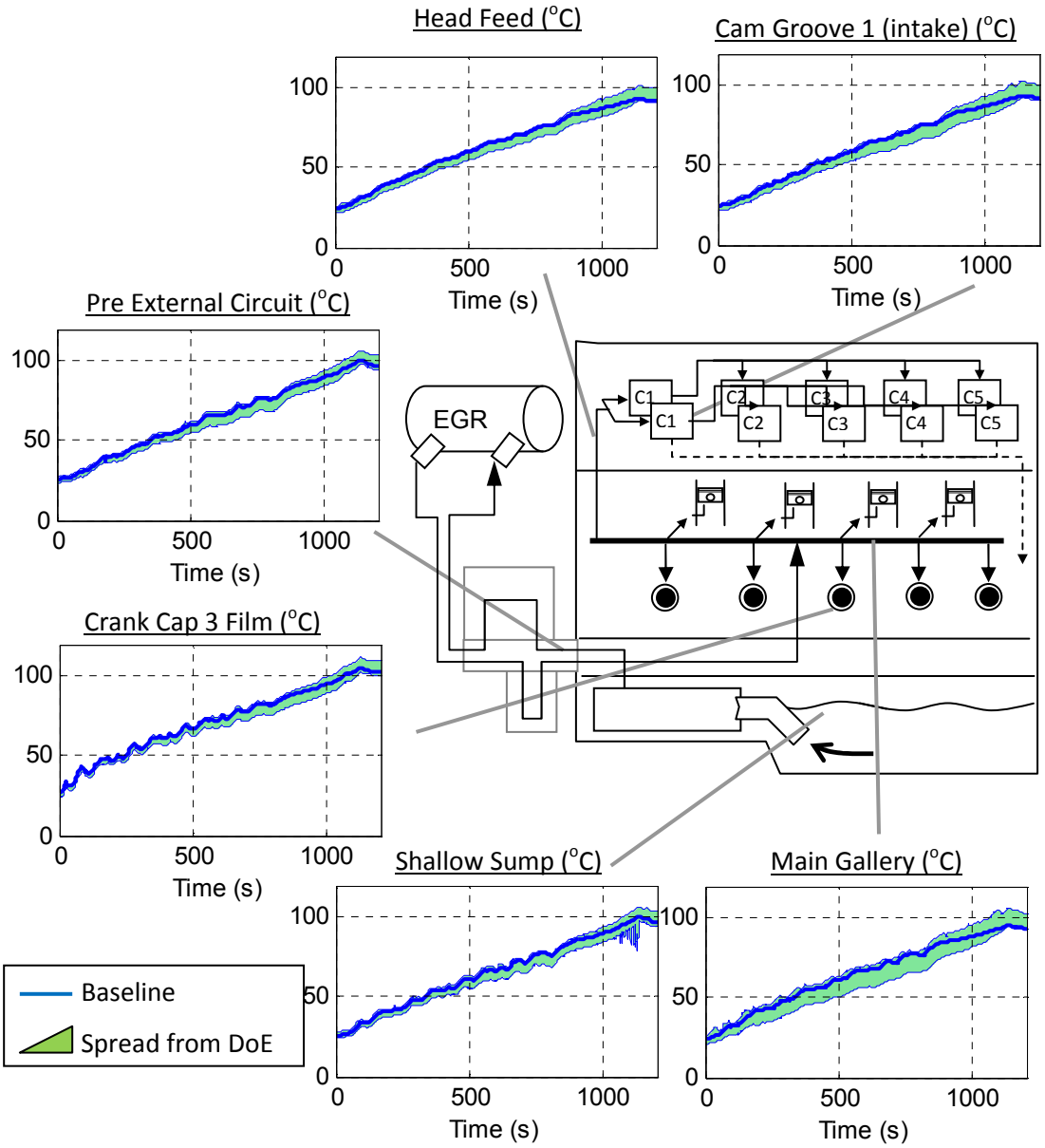


Figure 7.2: Selected oil temperatures for baseline setup and spread of results from DoE tests (All temperatures in °C)

[76] Because the DoE tests were all run with the VFOP operating at low flow, the *baseline* test in this case is not the production engine, but an engine running the VFOP in low flow setting but without the active thermal management devices.

[77] This line is the mean temperature rise for a series of 8 experiments which showed good repeatability.

Figure 7.3 shows selected coolant temperatures in the same way and two trends are apparent. At the front of the engine (Top hose, degas bottle, radiator etc.) the baseline temperature lies at the top of the spread from the DoE points. In the EGR cooler/oil cooler loop, the baseline coolant temperature lies in the middle or nearer the bottom of the spread of the DoE temperatures. As discussed in chapter 5, a portion of the coolant circuit at the front of the engine is isolated during warm-up which allows the rest of the circuit to warm-up faster.

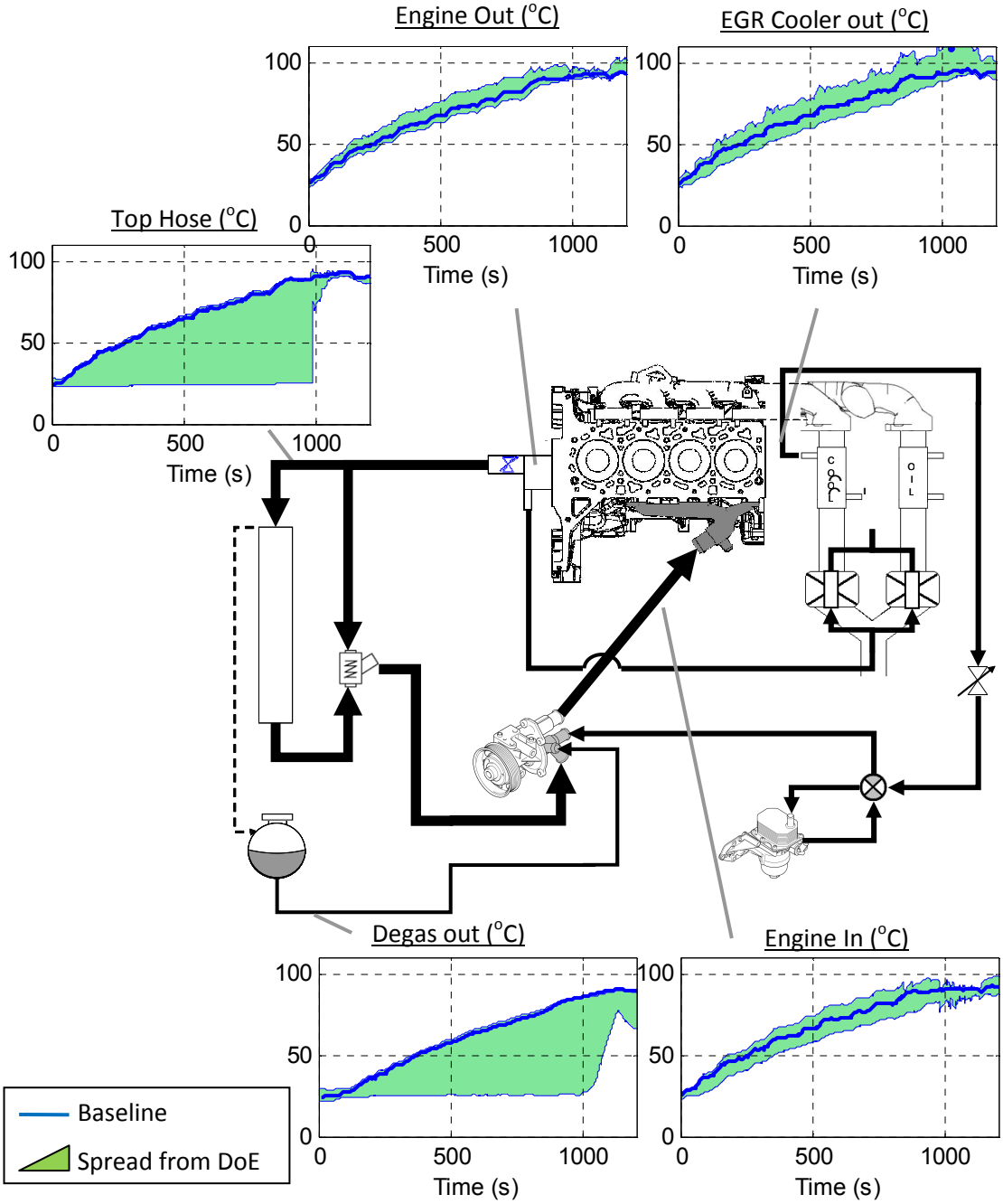
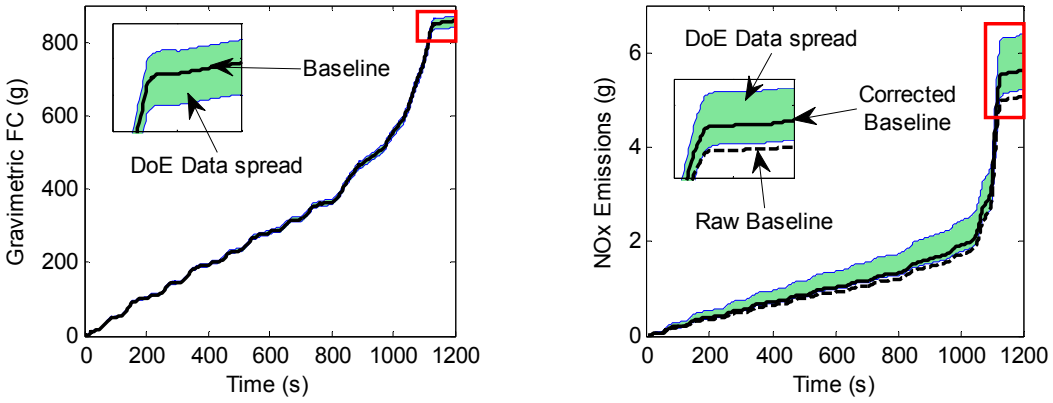


Figure 7.3: Selected coolant temperatures for baseline setup and spread of results from DoE tests (All temperatures in °C)

Figure 7.4 shows fuel consumption and NO_x emissions over the NEDC for the reference case and the DoE conditions. Despite the increased thermal inertia, the DoE results do offer some improvements in FC although there are clearly come conditions where fuel consumption was worse. It is difficult to draw any conclusions from the raw NO_x results as there appears to be a step change between the two data-sets: the raw baseline data is considerably lower than the DoE spread. Using the corrected NO_x results derived in chapter 3, there is a suggestion that NO_x levels for the *baseline* are approximately in the middle of the DoE scatter. This would be consistent with the fuel consumption results. Based on these plots the DoE had a larger impact on NO_x emissions than fuel consumption.



(a) *(b)*

Figure 7.4: (a) fuel consumption and (b) NO_x emissions for baseline setup and spread for DoE results (NO_x baseline results show raw data and corrected NO_x using relationship derived in chapter 3)

1.2.2. Response model analysis

(a) Overview

The response models can be split into two categories, although the building procedure and inputs are the same in both cases:

1. Models describing the thermal behaviour of the engine (temperature rise, heat flux…)
2. Models describing fuel consumption and emissions

The general model structure is shown in figure 7.5. The quantification of inputs and outputs depended on the particular metric and phase in the drive cycle and are detailed in table 7.3. All modelling was performed using the *MBC toolbox* in the Mathworks Matlab environment. The

model quality was assessed using R^2 for goodness of fit and adjusted R^2 and PRESS R^2 to check for over-fitting[78]. The modelling work was split between phases 1 and 2 of the NEDC to roughly split between the warm-up phase and a hot operating phase. In each case, a base model was produced with all linear terms and two-way interactions and a quadratic term was included for injection timing[79]. The model terms were then selected using PRESS and stepwise parameter selection using significant levels calculated from a t-test.

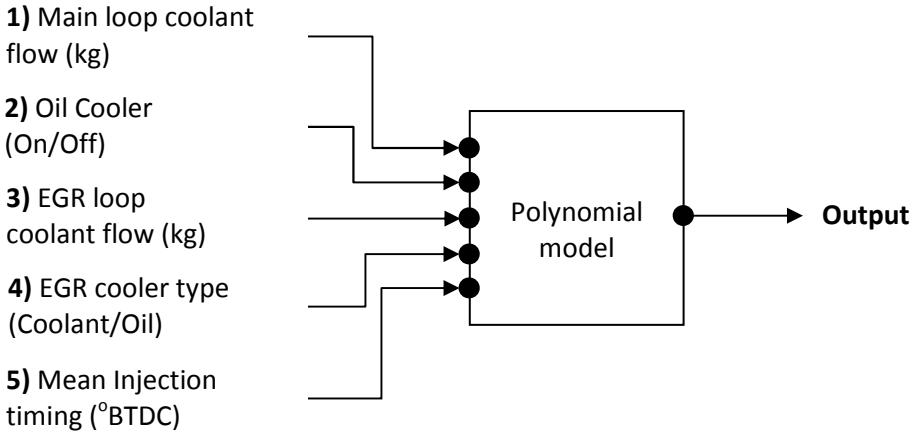


Figure 7.5: General model structure for describing engine behaviour over each phase of NEDC

INPUTS	
Metric	**Quantification**
Main loop coolant flow	Number of kg of flow over phase
Oil Cooler	Qualitative measure: On or Off
EGR loop coolant flow	Number of kg of flow over phase
EGR cooler type	Qualitative measure: Coolant or Oil
Mean injection timing	Average over test phase

OUTPUTS	
Metric	**Quantification**
Oil/coolant/metal temperatures	Temperature rise (phase 1) *or* Average temperature (phase 2)
Heat Flux	Total energy loss over test phase
Convective heat transfer coefficients	Average over test phase
Emissions/FC	Mass used/produced

Table 7.3: Variable quantification for modelling

[78] Although the models were mostly linear and there would not be a risk of over-fitting from a higher order polynomial, because the number of experiments was small compared to the number of inputs, there is a risk of over-fitting from the number of interaction terms.

[79] In some cases it was possible to use the data from DoE point 17 (see table 7.2) to fit quadratic responses for both main loop coolant flow and EGR loop coolant flow, however care needed to be taken due to the non-linear nature of the response.

(b) Category 1 Models: Warm-up and thermal modelling

Oil temperatures

The rise in oil main gallery temperature over phase 1 was dominated by three inputs which fitted the data well (R^2: 0.9, adj R^2: 0.9, PRESS R^2: 0.8). The oil warm-up rate was improved by engine-out coolant throttling, the use of oil cooler and the use of oil-cooled EGR. All three effects were of similar order of magnitude and a significant interaction term was measured between the oil cooler and EGR cooler. Figure 7.6 shows the response for oil cooler and EGR cooler: individually both controls offer about 6°C faster warm-up over phase 1. However when the two are used together, the overall benefit is significantly less than the sum of the two individual benefits (7°C).

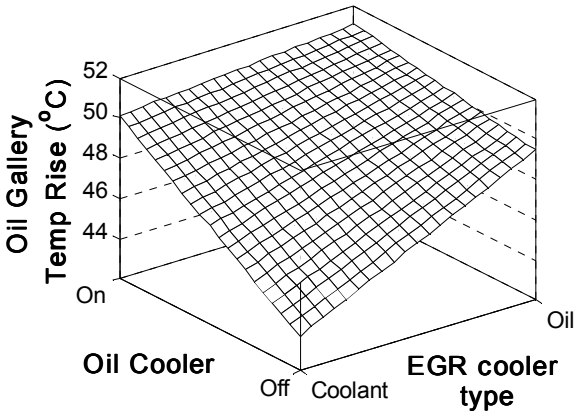


Figure 7.6: Effects of oil cooler and EGR cooler type on oil main gallery temperature rise

The majority of the oil circuit and bulk oil temperatures presented a similar behaviour with similar model qualities. These included oil temperature rises modelled at the crank shaft bearings, in the external oil circuit, the sump and the 1st cam shaft bearings. As shown in chapter 6, there is a significant correlation between oil supply temperature and the temperatures around the main circuit over phase 1 of the drive cycle. The only exception was the 5th cam shaft bearings which remained insensitive to the external oil circuit. Close inspection of the oil path in the cylinder head shows that the oil flow in bearing 1 is not only closer to the supply but also significantly larger as the flow for all 5 cam shaft bearings passes through this groove. The results suggest that bulk oil temperature is strongly affected by the external circuit, although the benefits are not seen in the far regions of the circuit with low flow. Engine coolant temperature appeared more influential in these regions. This is consistent with oil mass flow measurements from other authors on a similar engine [153]. They reported

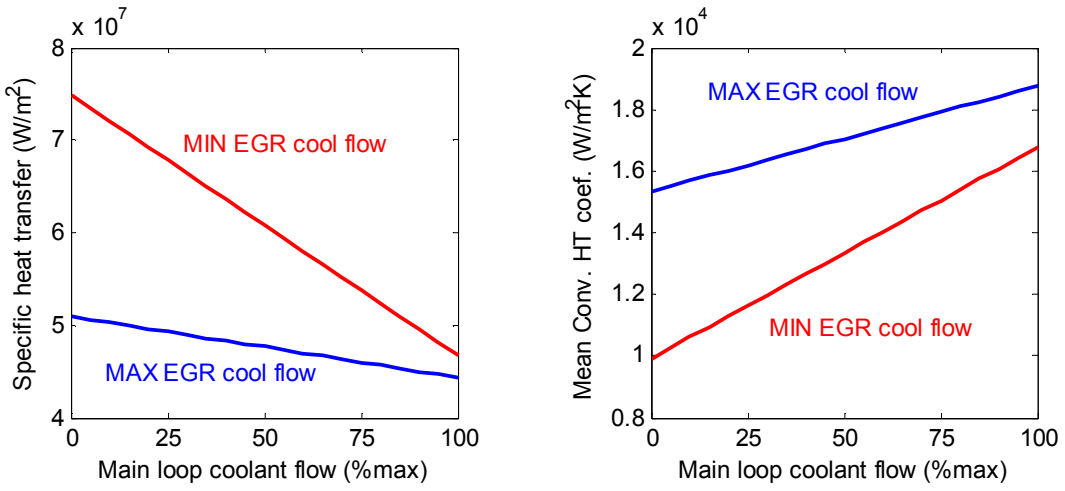
that only 30% of all oil flow reached the cylinder head which would then be split between each of the cam shaft bearings. In phase 2, the behaviour of oil remained similar to that in phase 1 because it does not reach its warmest temperature until the end of the drive cycle.

Coolant temperatures

When assessing the response models for coolant temperatures it was important to note the secondary effects of coolant throttling on engine inlet coolant temperatures. The rise in engine-out coolant temperature was modelled with similar goodness of fit to the oil temperature rise (R^2: 0.9, adj R^2: 0.9, PRESS R^2: 0.8). The same factors were influential in the response and throttling coolant flow in the main loop had the largest effect offering a 6^{o}C benefit over phase 1. The effects in the EGR loop were opposite to those for oil heating: bypassing the oil cooler and using coolant-cooled EGR were beneficial to coolant heating. Coolant temperature at engine inlet presented a more complex response showing significant interactions with coolant flow control.

Other coolant temperatures in the system are largely related to these two locations. Local temperatures in the coolant jacket are more or less related to engine inlet depending on their location (intake/exhaust side). There are a number of situations where both higher outlet and colder inlet temperatures are achieved yielding a significantly higher temperature difference over the engine. This has a knock-on effect on heat transfer in the coolant jacket and figure 7.7 shows the effect of coolant throttling on heat flux and convective heat transfer. Figure 7.7 (a) shows higher heat fluxes with lower coolant flows because of the stronger thermal gradient. However figure 7.7 (b) shows convective heat transfer coefficients are reduced at lower flow rates. This is explained by lower coolant velocities.

In phase 2 only the engine-out coolant throttle plays a significant role in coolant temperatures as the thermostat controls the evacuation of heat from the system. Later in this chapter it will be shown that under hot conditions, coolant flow throttling can increase coolant operating temperature.



(a) *(b)*

Figure 7.7: Effect of engine-out coolant flow on (a) heat flux and (b) convective heat transfer coefficient. The lower coolant flow causes increased heat flux because of a larger temperature gradient between the cylinder and colder inlet coolant; however convective heat transfer coefficient is reduced because of lower flow velocity.

Metal temperatures

Over phase 1 the results from the models for metal temperatures can be summarised in the following points:

- Cylinder head temperature (ECU control temperature) was strongly related to coolant flow in the main loop and coolant/oil heat exchange. The fastest warm-up is seen with the oil cooler off, coolant-cooled EGR and the main throttle closed. Other metal temperatures measured in the proximity of the valve train have a similar relationship and appear to be primarily linked to coolant temperature.
- Cylinder liner temperatures were also strongly linked to coolant temperatures and figure 7.8 shows the temperature response near the top of the cylinder liner to changes in oil cooler and EGR cooler control. This effect is opposite to that of main oil gallery temperature seen in figure 7.6. It should be noted that the effects near TDC were larger and presented more confidence than the results lower down the stroke, notable those below the coolant jacket.
- Crank cap temperatures linked to the oil film temperatures and responded to changes in oil temperature.

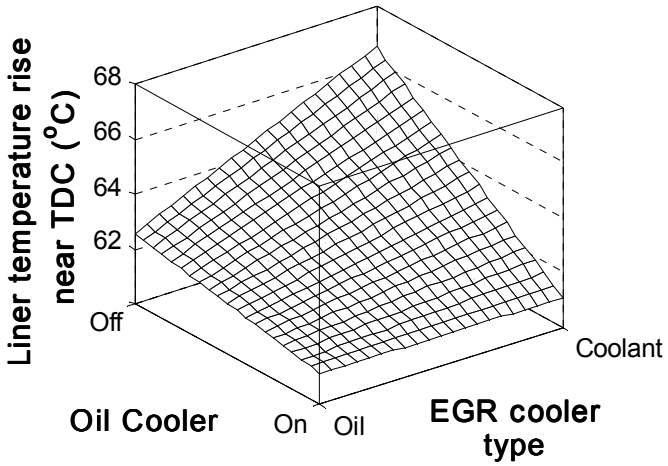
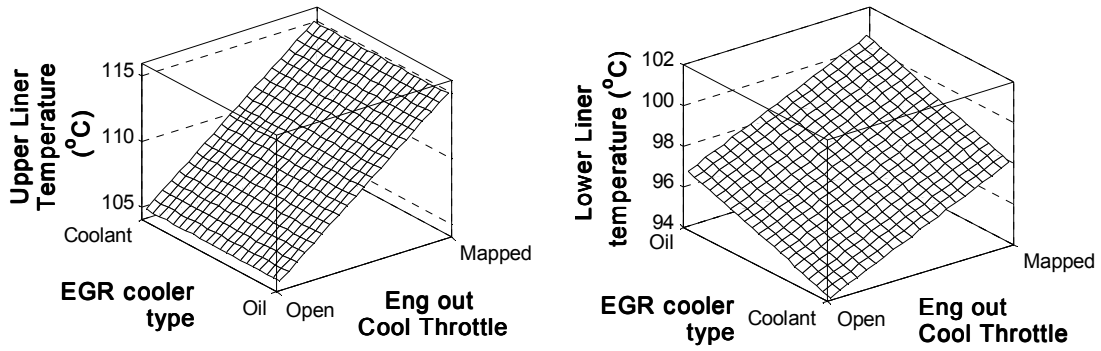


Figure 7.8: Effect of oil cooler and EGR cooler type on liner temperature near TDC

During phase 2 all liner temperatures were strongly affected by coolant flows. At the top of the liner, these effects were dominant compared to oil heating effects. Figure 7.9 (a) shows the predicted mean upper liner temperature over phase 2 with respect to engine-out coolant throttling and EGR cooler type. The coolant throttling effect was dominant and no effect from EGR cooler type was measured. However, below the coolant jacket the temperature was equally affected by the oil temperatures. Figure 7.9 (b) shows the impact on lower liner temperature showing effects of similar order of magnitude from both inputs. These results highlight the effects of piston cooling jets.



(a) *(b)*

Figure 7.9: Mean phase 2 temperature response for EGR type and main loop coolant flow at locations (a) near TDC and (b) near BDC

(c) Category 2 Models: Emissions and Fuel Consumption

The modelling work presented so far has demonstrated the control the flexible thermal management system offers during warm-up, however the benefits of these need to be assessed on a fuel consumption and emissions basis.

Fuel consumption and CO_2

Fuel consumption response models were fitted for each measurement technique available on the facility (carbon balance, gravimetric and ECU, see chapter 3). CO_2 results correlated strongly with carbon balance fuel consumption. Over the NEDC phase 1, the goodness fit was of lower quality than for the thermal modelling, with $R^2 \approx 0.8$, $adjR^2 \approx 0.7$ and PRESS $R^2 \approx 0.65$. This highlights that the fuel consumption measurement process is more susceptible to random error and disturbances. Because of these errors, some differences were apparent between the different models, however they were in agreement over the following points over phase 1:

- Throttling engine-out coolant flow during warm-up offered a 6g (2.3%) reduction in fuel consumption.
- Injection timing had a large effect on fuel consumption with a 3^o advance offering 20g (5.5%) benefit.

Over phase 2 of the drive cycle the models showed better fit characteristics ($R^2 \approx 0.95$, $adjR^2 \approx 0.94$ and PRESS $R^2 \approx 0.92$). Again there were some differences between the models from each of the measurements, but they each agree on the following points:

- Injection timing has the largest effect on fuel consumption at this stage (3^o advance giving 2.3% benefit).
- Using oil-cooled EGR and oil cooler improved fuel consumption (0.5%).

The large impact of injection timing on fuel consumption shows that the DoE is ill-conditioned for assessing benefits during phase 2 and that a smaller range should have been used. As a result, some effects of other variables may be masked amongst the response. This may also explain why such a good level of fit was achieved with these models.

NO_x emissions

Over phase 1 a good level of fit (R^2: 0.85, adj R^2: 0.82, PRESS R^2: 0.78) was achieved with a simple model. The response was again dominated by injection timing with a 3^o advance increasing NO_x emissions 30% (0.45g). EGR cooler type was the second largest effect with oil-cooled EGR reducing NO_x by approximately 6%. There is also a small interaction term shown in figure 7.10 which reduces the effect of injection timing to 0.3g (20%) when main engine-out throttle is closed.

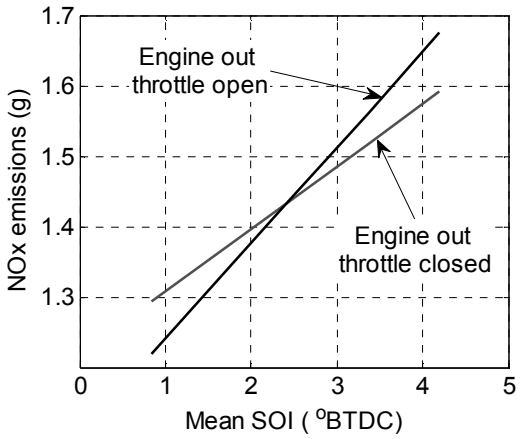


Figure 7.10: Effect of injection timing on NO_x emissions under different engine-out coolant throttle calibrations

The effect of injection timing is well known. The interaction term between engine-out coolant throttle and injection timing could result from the strong link to upper engine temperature (see category 1 models). The changes in temperature around the combustion chamber during warm-up would affect ignition delay and subsequent premix combustion.

The effect of the EGR cooler type is less obvious. Figure 7.4 (c) showed that oil-cooled EGR was more effective over cold-start NEDC which would reduce intake gas temperature and combustion temperature. It is thought that this is primarily due to the oil being colder than the coolant for the majority of the cold-start drive cycle, however, close inspection on the final 100 seconds where oil is hotter than coolant show that there is still no switch in temperatures, and oil EGR cooling is always more effective. The work published by Torregrosa *et al.* [89] estimated reductions in NO_x emissions of approximately 1% per 1^oC reduction in inlet manifold temperature. Over phase 1 inlet manifold temperature is on average5^oC colder with oil-cooled EGR, so the 6% reductions agrees well with the published data.

Over phase 2 despite a lower model quality, injection timing continues to dominate the response (3^{o} advance causing 8% rise). EGR cooler type also had a large effect with switching from coolant to oil-cooled gases reducing NO_x by 5%. As inlet manifold temperatures were on average $12^{o}C$ colder, this reduction in NO_x is lower than expected based on the comparison with the published data. Further analysis of the work by Torregrosa *et al.* [89] suggests the effects reduce under higher load conditions. This would then agree with results over phase 2.

The model showed reduced effectiveness of the oil-cooled EGR when used in conjunction with the oil cooler. Higher reductions in NO_x are achieved with the oil cooler off as shown in figure 7.11. These interactions can be explained by the higher oil temperatures achieved during cold-start NEDC because of the oil heating from the coolant.

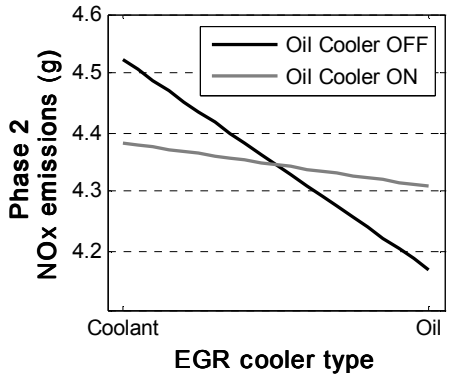
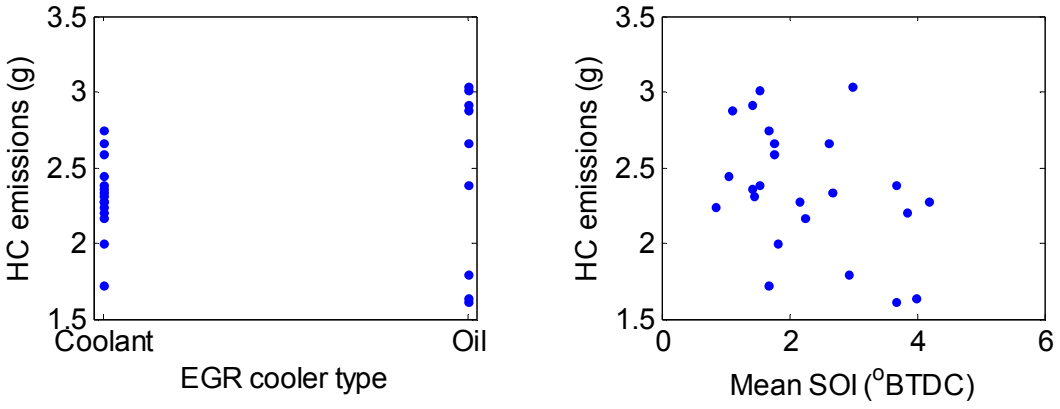


Figure 7.11: Effect of EGR cooler type on NO_x emissions with oil cooler on or off

CO and HC emissions

Good quality response models were produced for CO over both phases 1 and 2 (R^2: 0.9, adj R^2: 0.89, PRESS R^2: 0.86). Over phase 1, CO emissions could be reduced 5% through engine-out coolant throttling and faster warm-up but would be increased 11% when using oil-cooled EGR (Effectively this is the opposite effect to NO_x). As for NO_x emissions, this is in good agreement with the work published by Torregrosa *et al.* [89]. Over phase 2 the largest effects were again caused by changes in injection timing and EGR cooler type (approx. 14% increase for 3^{o} retard or switch to oil-cooled EGR). Engine-out coolant throttling continued to have an effect with mapped flow offering 11% reduction. It is important to note over phase 2 that tailpipe CO emissions are low and usually of less interest because catalytic convertors are very effective at these operating conditions.

It was not possible to fit appropriate response models to the HC results over either of the NEDC phases without over fitting the data[80]. As a result, the raw data was the only source of information, which is often difficult in designed experiments due to the simultaneous variations of inputs. HC emissions over phase 1 are plotted in figure 7.12 against EGR cooler type and injection timing, however no clear trends are apparent. These results would be expected to follow similar trends to CO emissions.



(a) *(b)*

Figure 7.12: Trends in phase 1 HC emissions for(a) EGR cooler type and (b) injection timing

1.3. Modelling results discussion

It is clear that the active thermal management had a significant effect on the warm-up behaviour of the engine with temperatures at many locations varying by more than 10°C over the design space. The active control of heat flows during warm-up has shown that throttling overall coolant flow contributes to faster warm-up throughout the engine. During phase 2 it can also be used to allow the coolant to operate at a higher average temperature which will continue to promote warm-up by virtue of a higher thermal gradient. The response models have clearly identified that the EGR cooler and oil cooler calibrations impact on heat flows during warm-up either to coolant, cylinder head and upper *or* to oil, main bearings and lower liner. As these devices did not significantly change the available heat energy, a trade-off appeared between ***upper*** and ***lower*** engine warm-up: this trade-off, generated from the response models is illustrated by figure 7.13.

[80] A larger number of coefficients could increase the R^2 value, but gave critically low adjusted and PRESS R^2 values.

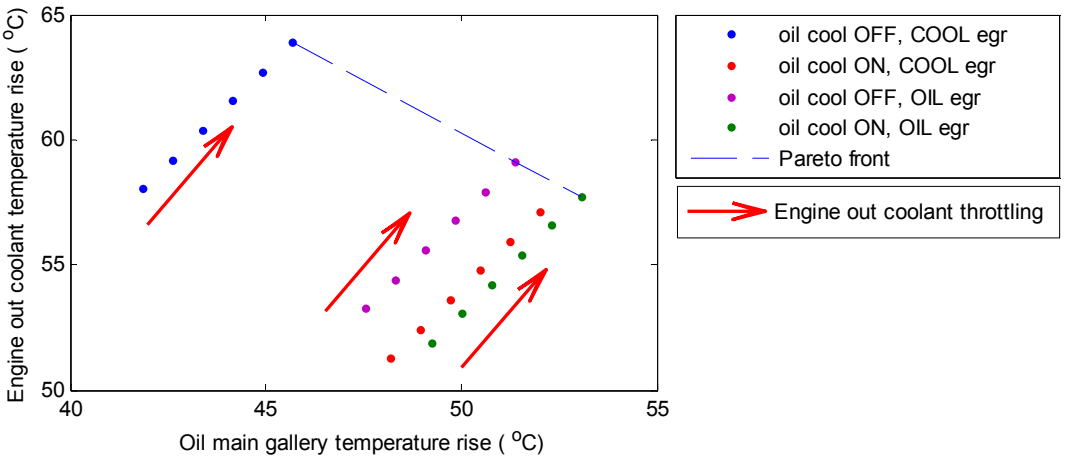


Figure 7.13: Coolant and Oil and upper engine warm-up trade-off resulting from oil cooler and EGR cooler calibration

The thermal trade-off appears to have an impact on NO_x and CO emissions through its effect on EGR gas cooling. A similar trade-off between these emissions species could therefore be imagined in response to the oil cooler and EGR cooler control. These two parameters had much smaller effects on fuel consumption and CO_2 emissions showing there should be scope for performance benefits through optimisation.

1.4. System Optimisation

1.4.1. Optimised calibrations

The engine in this study tends to be limited by NO_x limits rather than smoke or HC emissions because of the vehicle application was a light duty truck which works the engine harder relative to its power output. As a simple approach, the active thermal management will be optimised based on a NO_x-FC trade-off: effectively minimising fuel consumption within a NO_x constraint. Because the relationships between the inputs and outputs are simple in this modelling approach, optimisation is possible by visual inspection, but similar results were also obtained using commercially available optimisation software such as Mathworks *MBC toolbox Calibration Generator (CAGE)*. The optimisation process was carried out separately for phases 1 and 2 and the resulting calibrations are presented in table 7.4.

Variable	Phase 1	Phase 2
Engine-out throttle	Mapped	Mapped
EGR loop coolant throttle	Min flow	Min flow
EGR cooler type	Oil-cooled	Oil-cooled
Oil cooler	On	Bypass
SOI	Advance 1.2^{o}	Advance 1.5^{o}
Anticipated FC benefit at iso-NO_x	12g (-3.2%)	10g (-2%)

Table 7.4: Optimised setting for controlled thermal management system and expected benefits from response model

Before testing these calibrations it is worth discussing the settings of the control variables in light of the modelling work. As previously discussed, phase 1 represents the warm-up phase and the optimised setup relies on targeting fast warm-up and channelling heat to the engine oil. During phase 2, the changes in oil viscosity become much smaller due to the exponential relationship with temperature. Also, phase 2 is the main contributor to NO_x emissions because the engine is both hotter and operating under higher loads. The combination of these factors means that the effects of improved charge air cooling by colder oil is much more significant than benefits in fuel consumption from hotter oil. Consequently, the optimised setup aims to keep EGR gas temperatures to a minimum.

Considering first the setup for phase 1:

- Mapped engine-out flow and minimum EGR loop flow aim to maximise engine warm-up rate;
- The use of oil cooler and oil-cooled EGR aim to maximise oil warm-up rate;
- Advanced SOI counteracts the ECU tendency to retard fuel injection with faster warm-up rates, as demonstrated in chapter 5.

For phase 2,

- The mapped engine-out coolant throttle aims to raise the target coolant temperature at part load.
- The combination of oil-cooled EGR and bypassed oil cooler aims to provide maximum EGR gas cooling. As oil temperatures remain lower than coolant temperatures during the majority of phase 2, bypassing the oil cooler keeps the oil temperatures lower.
- Advanced SOI aims to trade-off the benefits in NO_x from improved EGR gas cooling to yield benefits in fuel consumption.

Although two separate optimised setups have been proposed aiming for minimum fuel consumption at iso-NO_x for each of the two drive cycle phases, clearly a combined optimum setup would also be desirable. The optimisation approach for phase 2 assumed a similar calibration over the whole drive cycle. This will clearly not be the case in a calibration optimised for both phases where a switch occurs between phases. As a result the validation test is unlikely to meet the model prediction because of discrepancies of the engine thermal state at the start of phase 2 between this test and the model training data.

1.4.2. Validation

Based on the experimental facility capabilities established in chapter 3, each of the conditions were run for the minimum number of times to achieve significant differences at 95%. As some time had passed between the *DoE tests* and the *validation tests* with some small disturbances to the engine[81], a series of tests in a *baseline* calibration were conducted to better quantify the benefits from the optimised setups.

Figure 7.14 compares phase 1 FC and NO_x predictions to measured results from the validation testing. The fuel consumption results agree well with the model prediction, however there are some inconsistencies with the NO_x results. The offset of NO_x is observed in both the baseline and the optimised setup and the relative differences between the optimised and baseline setups are similar for both model and testing. If the empirical NO_x correction factor derived in chapter 3 is applied to each of the test results, the offset is no longer apparent.

Table 7.5 compares other measures from the validation testing and the modelling predictions. The selected temperature rises were chosen to represent the upper and lower parts of the engine. There are small differences of about 2^{o}C between the model predictions and the actual engine behaviour throughout but the relative differences between baseline and optimised calibrations are consistent. Fuel consumption and NO_x emission results are also detailed for completeness, and there is good agreement between the carbon balance and the gravimetric results.

[81] To avoid the modelling work delaying experimental work, the experiments with the exhaust gas heat exchanger were performed between the DoE testing and the validation testing. Although care was taken to return the engine to the same setup as for the DoE, there have inevitably been changes in connectors and some pipes. Also, the DoE tests were performed in December, during a particularly cold period whereas the validation tests were performed in June in more mild conditions; this has a significant effect of NO_x emissions because of large changes in ambient humidity.

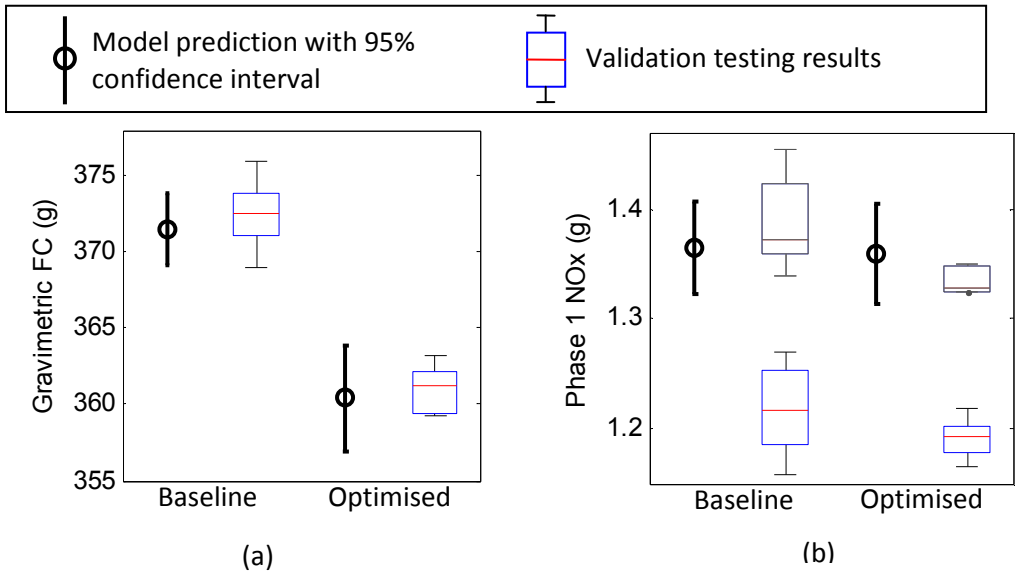


Figure 7.14: Comparison of predicted and measured (a) fuel consumption and (b) NO_x showing model prediction confidence and validation measurement spread for phase 1 optimisation (grey box plots show corrected NO_x results)

	Variable	Baseline		Optimised phase 1	
		Predicted	Measured	Predicted	Measured
Temperature (°C)	ECU	53	55	59	61
	Oil Main Gallery	49	52	53	55
	Crank cap oil film	49	51	53	54
	Liner Mid stroke	55	57	59	61
FC/ Emissions (g)	Carbon Balance	371	373	366	363
	Gravimetric	372	372	360	361
	NO_x	1.37	1.22	1.36	1.19
	Corrected NOx	1.36	1.37	1.36	1.34

Table 7.5: Comparison of predicted and actual behaviour for phase 1 for selected temperature rises, fuel consumption and NO_x emissions

Figure 7.15 and table 7.6 shows the same results for the optimisation during phase 2. Again there is good agreement for the thermal behaviour. The fuel consumption results also give good agreement and whilst there the agreement between predicted and measured NO_x is larger than for phase 1, these are still acceptable. The lower quality results for NO_x over this phase are most probably a result of the sensitivities of NO_x emissions in each phase of the

drive cycle. Phase 2 represents over 70% of total NEDC NO_x and small changes in engine control parameters will have significant effects.

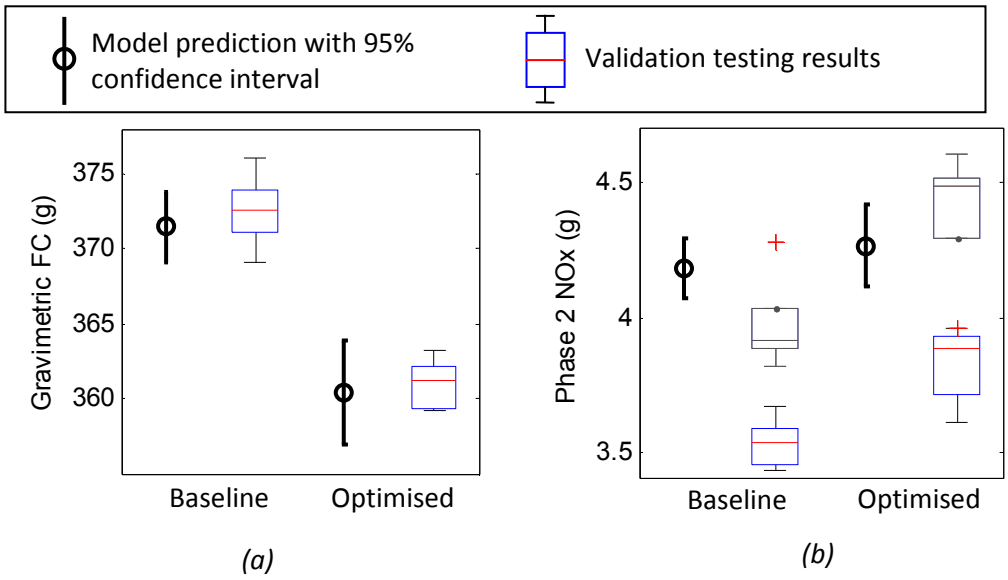


Figure 7.15: Comparison of predicted and measured (a) fuel consumption and (b) NO_x showing model prediction confidence and validation measurement spread for phase 2 optimisation (grey box plots show corrected NO_x results)

	Variable	Baseline		Optimised phase 2	
		Predicted	Measured	Predicted	Measured
Temperature (°C)	ECU	92	93	103	103
	Oil Main Gallery	85	87	91	94
	Crank cap oil film	90	92	95	98
	Liner Mid stroke	93	94	99	101
FC/ Emissions (g)	Carbon Balance	506	517	496	504
	Gravimetric	504	505	494	497
	NO_x	4.18	3.48	4.3	3.92
	Corrected NO_x	4.18	3.93	4.3	4.52

Table 7.6: Comparison of predicted and actual behaviour for phase 2 for selected temperature rises, fuel consumption and NO_x emissions

2. *System behaviour under fully warm conditions*

The main focus of this thesis is to optimise the warm-up behaviour. However, there was also potential for using the system under fully warm conditions to improve system performance. Although a full DoE approach was not undertaken, a series of experiments were conducted aimed at assessing two aspects:

1. Benefits from engine-out coolant flow mapping under hot conditions.
2. Behaviour of oil-cooled EGR under hot conditions.

2.1. Effect of engine-out coolant throttle

To assess the effect of coolant flow mapping under hot conditions a series of hot-start NEDC tests were conducted with and without engine-out throttle control. Measured coolant flows are shown in figure 7.16 along with the impact on selected temperatures. The open loop control had been calibrated to give a near constant cylinder-head temperature by adjusting coolant throttle position based on engine speed and fuelling quantity. This approach is very similar to that presented by Heiduk *et al.* [13].

For most of the NEDC the system almost supresses coolant flow but even during the final 100 seconds flow is about half that of the non-controlled condition. This effectively increased coolant temperatures by about 10^{o}C during the low power phase, but allowed the coolant temperature to drop to the baseline levels at higher powers. Cylinder liner temperature correlated with the coolant temperature due to the strong interactions with the coolant jacket. Oil temperatures were increased but by slightly less and were approximately 8^{o}C hotter throughout the cycle.

The impact of these thermal changes on fuel consumption and emissions is summarised in figure 7.17. Mapping the engine-out coolant flow gave a 1% reduction in fuel consumption over the NEDC but incurred a 5.5% NO_x penalty. Feed-gas CO emissions were also reduced by 12% although under fully warm conditions these are of less interest. These results are explained by the higher temperature offering reduced frictional losses but increased combustion temperatures. Statistical analysis shows that all of these differences were significant at 95% confidence level.

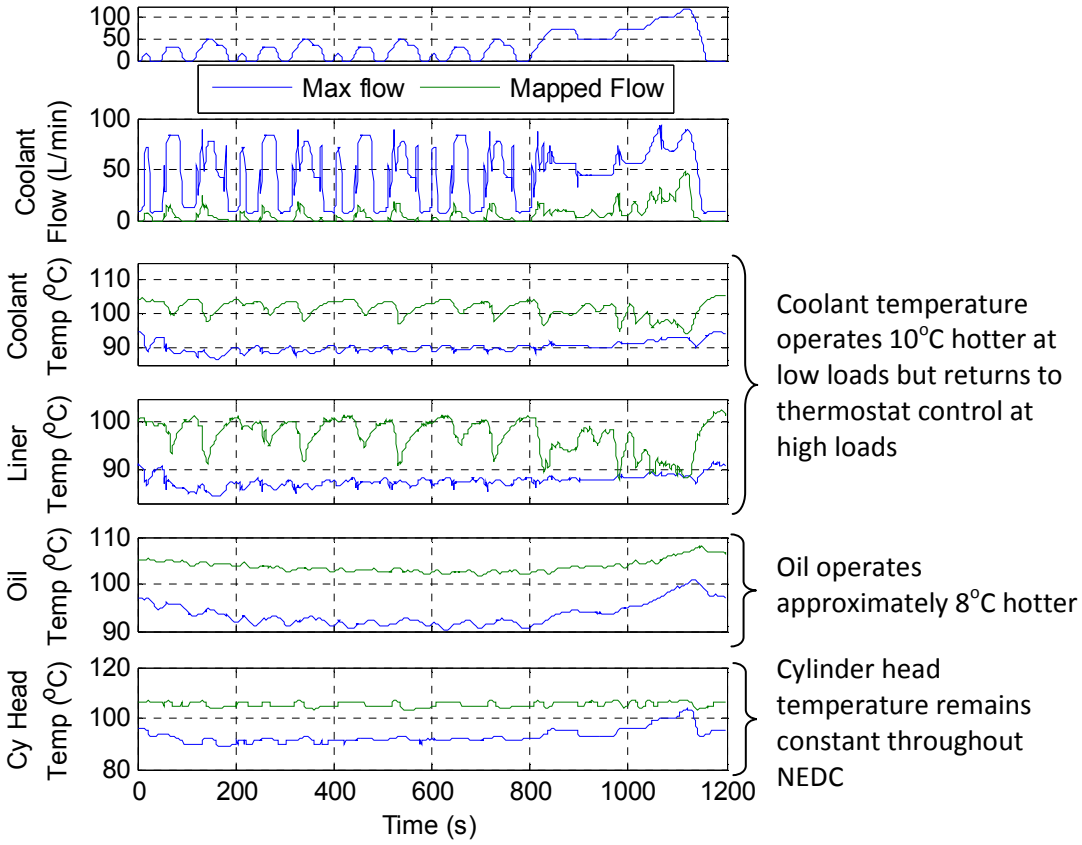


Figure 7.16: Impact of engine-out throttle mapping during hot NEDC on coolant, liner, oil and cylinder head temperatures

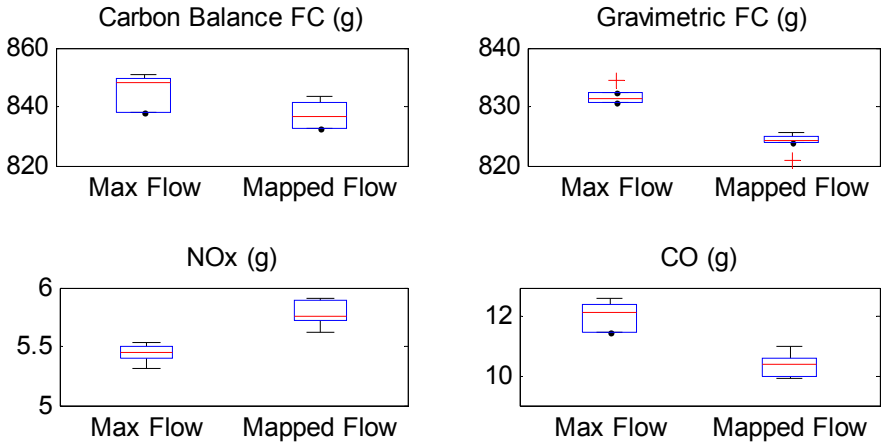


Figure 7.17: Hot-start NEDC fuel consumption, NO_x and CO emissions for max and mapped engine-out coolant flow

Using the system in way results in similar behaviour to that presented by a number of authors [22, 23, 37] and discussed in chapter 2. In each of these cases the authors attempted to increase warm-up rate through flow throttling but only managed to change steady state coolant operating temperature. Steady state results for fully warm conditions presented by

Brace *et al.* [23] showed a 2% reduction and 12% increase in NO_x when temperatures were raised 15°C by flow throttling. The changes measured in this chapter are similar to these changes. In their work, Brace *et al.* did not see any benefits from cold start drive cycle but did not perform any hot-start tests. This work demonstrates that the steady state benefits in FC do translate to hot-start drive cycle benefits.

2.2. Effect of oil-cooled EGR

Oil-cooled EGR offered significant NO_x benefits over a cold-start NEDC because it was more effective at cooling EGR gases. Under cold conditions the benefit can be in part explained by oil temperatures being colder than coolant temperatures throughout the warm-up period. If the oil-cooled EGR could offer similar or improved behaviour than coolant-cooled EGR under hot conditions then a production hardware setup could be based around a single oil-cooled EGR system. This would obviously be more attractive from a cost perspective than the prototype dual system.

A series of hot-start NEDCs were conducted switching between coolant-cooled and oil-cooled EGR. In both cases the engine-out coolant throttle was mapped and the oil cooler was used. This was important to maintain a sensible operating temperature for the oil despite the additional heat input from the EGR gases. Figure 7.18 shows the impact on coolant, oil and gas temperatures over a hot-start NEDC. Firstly it is clear that there is no impact on coolant temperatures as the thermostat ultimately controls this settling temperature, although coolant throttling has been shown to influence this. On the other hand, the oil temperature in the main gallery is approximately 5°C hotter throughout the test when using oil-cooled EGR. The effect on EGR and intake manifold gases is a little more complex:

- Oil-cooled EGR gas cooling appears to be more effective, with larger differences at higher loads.
- The effects are more apparent in the inlet manifold at lower loads where EGR rates are higher. The colder EGR gases at higher loads have less impact because the EGR gas flow is small relative to the fresh air flow.

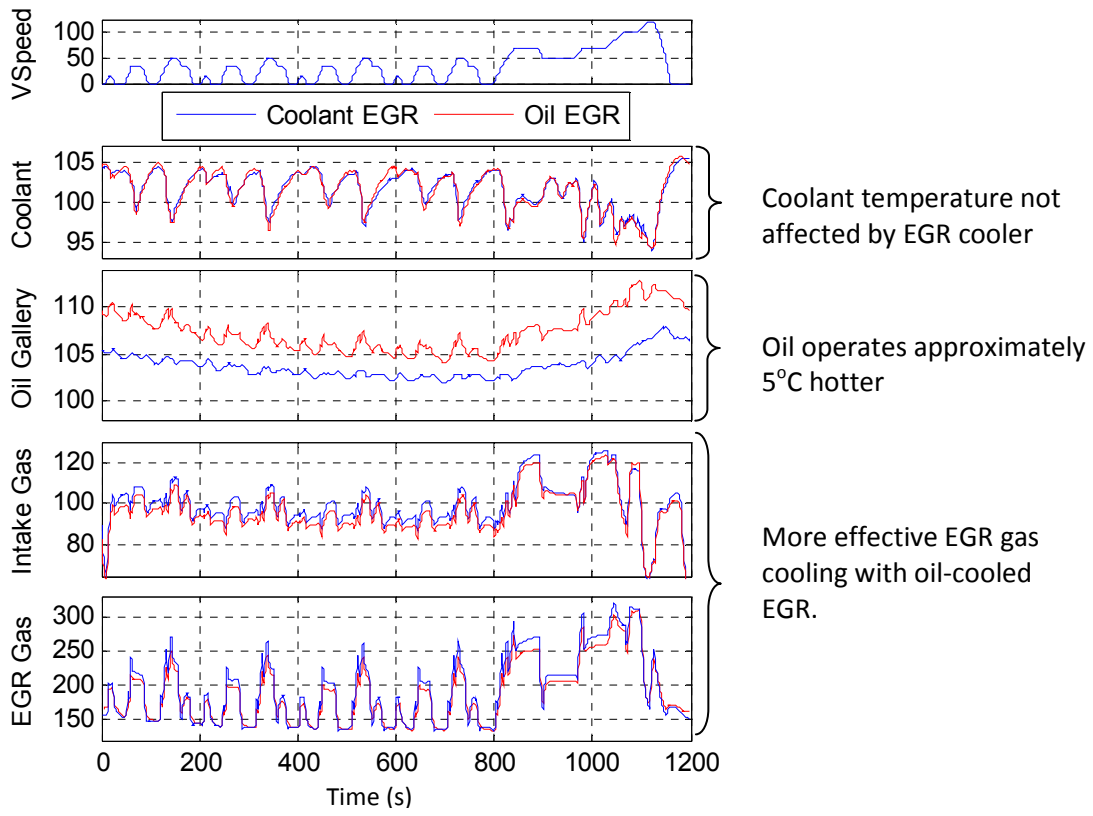


Figure 7.18: Effect of EGR cooler type on coolant, oil and gas temperatures during hot-start NEDC (all temperatures in °C)

As the oil-cooled EGR offered benefits in terms of EGR gas cooling over coolant-cooled EGR, it was thought this may be improved further by varying oil flow rate using the variable flow oil pump. Again a number of hot-start NEDC tests were conducted at three different flow settings[82]. Detailed analysis of the temperatures in each of the calibrations showed that the higher oil flow rates did not improve EGR gas cooling and in fact were less effective than the lower oil flow. This is explained by referring to the oil temperature at pump delivery presented in chapter 6: higher oil flow rates resulted in higher oil temperature because of the additional pumping work. The corresponding effects on fuel consumption and emissions over the drive cycle are summarised in figure 7.19:

- The difference between calibrations 1 and 2 is the switch to and from coolant- to oil-cooled EGR. This results in a rise in fuel consumption (1%) and CO (4%), but a drop in NO_x emissions (0.5%). The change in fuel consumption is statistically significant, however the relatively small changes observed in NO_x and CO were not significant at 95% confidence level.

[82] The three settings where the same as those used in chapter 6: minimum flow with target gallery pressure 1-2bar; flat 3bar target pressure and maximum flow.

- Calibrations 2, 3 and 4 represent variations in oil flow rate whilst using oil-cooled EGR. Between the minimum and maximum oil flow there is an increase in both fuel consumption (1.8%) and NO_x (1.6%), but a reduction in CO emissions (2%). The increased fuel consumption is a result of higher oil pumping work described in chapter 6. The increased NO_x emissions are a result of the less effective EGR gas cooling. The difference between the 3bar flat map and the maximum flow calibrations is very small because under hot conditions the maximum oil flow is almost equivalent to a 3bar oil supply.

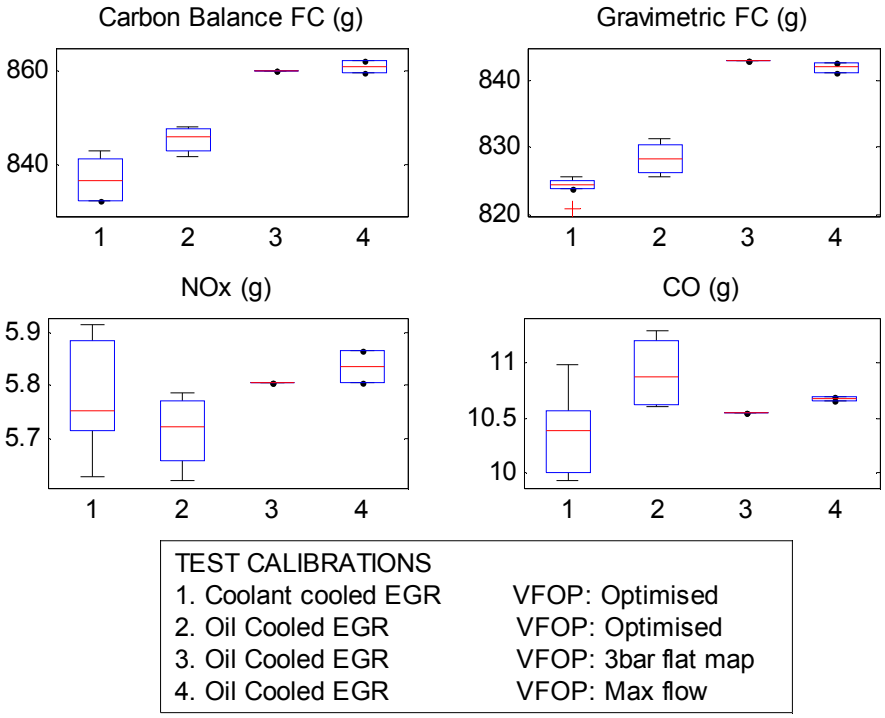


Figure 7.19: Hot-start NEDC fuel consumption, NO_x and CO emissions for oil-cooled EGR at varying oil flow rates

The results from the oil-cooled EGR showed no statistically significant reduction in NO_x under hot conditions. However, there was also no significant increase, showing that the oil-cooled EGR system can function appropriately under hot conditions. The results from variations in oil flow rate have highlighted that under hot conditions, while oil-cooled EGR can perform at a similar level to coolant-cooled EGR, this is only possible if used in conjunction with a variable flow oil device. The results presented in this chapter are only valid in the context of NEDC duty cycle, however under higher load conditions EGR is not often used and there should not be significant issues for oil temperature control.

3. *Chapter summary and conclusions*

The first section of this chapter has described a complete DoE calibration task applied to the engine thermal management system. The experiment consisted of 17 test points and good control and repeatability was demonstrated. Polynomial based statistical models were built to represent both thermal and macroscopic behaviour over each phase of the NEDC. These models were subsequently used to optimise the engine warm-up calibration. The second part of this chapter considered the viability of the system optimised for engine warm-up to perform under fully warm conditions. The following points may be concluded from this chapter:

1) Other than for HC emissions, the modelling approach used in this chapter provided useful descriptions of the system behaviour. It is useful to split the modelling into phase 1 and phase 2 of the drive cycle because of the relative importance of each phase for different outputs. Phase 1 provides a good characterisation of engine warm-up.

2) As observed in chapter 5, coolant flow throttling in the main external loop improved engine warm-up by isolating parts of the circuit and effectively reducing the overall thermal inertia. Local temperature measurements showed this improved warm-up in the entire engine, with temperatures approximately 6^{o}C hotter at the end of NEDC phase 1. This reduced FC and CO by 2% and 5% respectively over NEDC.

3) Variations in the external oil circuit (oil cooler and dual EGR system) allow the modulation of heat flows during warm-up between oil and coolant. Since this does not significantly change the amount of heat energy available, a trade-off appeared between upper-engine warm-up (controlled by coolant) and lower-engine warm-up (controlled by oil). Variations of 8^{o}C and 6^{o}C were seen at the end of phase 1 for oil and coolant temperature respectively.

4) NO_x emissions could be reduced 5-6% throughout the drive cycle by using oil-cooled EGR but CO emissions increased by 11%. This effect was explained by more effective EGR gas cooling, with EGR gas temperatures up to 25^{o}C colder than with coolant-cooled EGR.

5) Optimised calibrations were predicted to give FC benefits of 3.2% and 2% over phase 1 and a 2 respectively under iso-NO_x conditions. Validation testing confirmed results for both phases showing that the combined DoE and NEDC optimisation approach does give meaningful results.

6) The system demonstrated adequate performance under fully warm conditions. Engine-out coolant flow control allowed the engine to operate 8-10^{o}C hotter under part load conditions giving 1% FC benefit but 5.5% NO_x penalty. Oil-cooled EGR did not have adverse effects on EGR gas cooling, provided it is implemented in conjunction with a variable flow oil pump.

Chapter 8 - *Improving the engine warm-up trade-off: Exhaust gas heat exchanger*

In the previous chapter a number of models were developed based on the DoE results to assess the thermal and macroscopic behaviour of the engine. In the modelling of warm-up rates at different locations in the engine, a trade-off was observed between warm-up of the coolant, cylinder liners, engine head or ***upper engine*** *and the oil and bearing temperatures or* ***lower engine****. This trade-off was a result of the limited energy available during warm-up and that while the system could control energy flows, it did not increase the available energy per unit thermal mass*[83]*.*

In the literature review in chapter 2, two methods for improving engine warm-up were identified: the first through reduced thermal inertia has been implemented through coolant flow throttling. The second involves external heat addition and will be implemented in this chapter. An exhaust gas heat exchanger was sourced and integrated into the system on the dynamometer whilst considering real world, on-vehicle constraints. The aim was to use otherwise waste heat to enhance the warm-up rate beyond what has been achieved in the previous chapter. Although this work was not originally included in the project plan, it has been included following the results from the previous chapter.

[83] There will be small changes in energy as a result of the small changes in fuel consumption, however these will not be significant in terms of affecting the warm-up rate.

1. *Heat exchanger hardware and installation*

1.1. Heat exchanger hardware

A suitable heat exchanger for the exhaust pipe was sourced. The device was taken from a production, medium-sized passenger car, and in that application it was used to improve cabin heating performance from cold-start. The exhaust gas heat exchanger layout is shown schematically in figure 8.1 and a photograph is shown in figure 8.2.

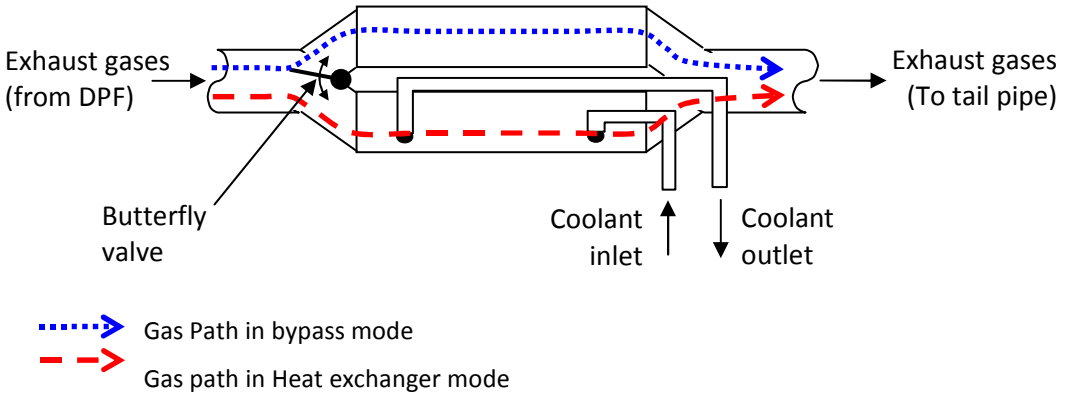


Figure 8.1: Exhaust gas heat exchanger layout and operation

Figure 8.2: Exhaust gas heat exchanger

The heat exchanger itself is a simple gas to liquid heat exchanger although some complexity arises from its installation within the engine. The primary purpose of the heat exchanger is to allow some waste heat from the exhaust to be captured by the coolant to assist with warm-up. Once the engine is warm the heat exchanger is unnecessary and indeed counter-productive. In order to avoid undesirable heat gain which must then be rejected by the radiator and also to avoid boiling the coolant flowing through the heat exchanger when the engine has warmed up, it is necessary to divert the gas flow. Diverting or stopping the coolant flow will always leave

stagnant coolant within the heat exchanger which will boil as hot exhaust gases pass through the device. To this end, the heat exchanger is equipped with an integral bypass leg and vacuum operated actuator controlling a butterfly valve. A simple control algorithm was put in place which switched the gas flow to bypass mode when a critical coolant temperature exiting the heat exchanger was reached. For practical uses of the heat exchanger this would be set between 90-100^{o}C, however the device could also be disabled by setting a low coolant temperature threshold.

1.2. Heat exchanger installation

A number of practical aspects were taken into account when installing the exhaust gas heat exchanger both on the gas side and the liquid side. It was decided that a near production setup should be adopted, within the constraints of the experimental facility. Each aspect will now be discussed separately.

1.2.1. Gas side installation

The gas side installation of the exhaust gas heat exchanger offers less flexibility than the liquid side. On a production vehicle expected to meet EURO VI emissions standards, it is almost certain that the exhaust pipes will require a catalyst and particulate filter (DPF)[84]. Because of the regeneration requirements of the DPF and the light off requirements of the catalyst, the exhaust gas heat exchanger would need to be installed downstream of these devices. The experimental facilities used in this experiment did not include a DPF because of issues of regeneration and filter state that could affect the repeatability of the measurements. However to achieve a similar situation to the production engine, the exhaust gas heat exchanger was installed several meters downstream from the catalyst. Not only would this affect gas side temperature at the heat exchanger, but would also give realistic hose lengths for the liquid side. A photograph of the gas side installation is shown in figure 8.3.

Insulation of the exhaust pipe was not considered in this work, however this may become important in getting heat to the exhaust components. The limitations will be whether heat is rejected sufficiently without damage to the after-treatment devices at fully warm operating temperatures.

[84] At the SIA international Diesel conference 2010, PSA announced all future development would be based around a selective reduction catalyst and particulate filter configuration.

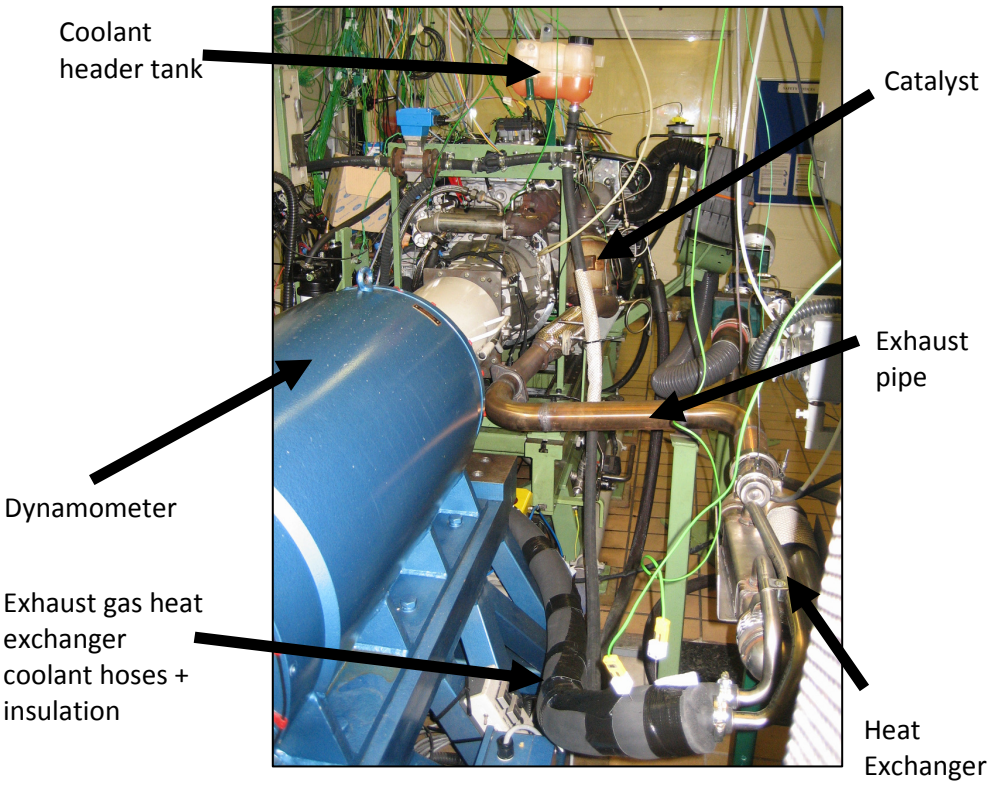


Figure 8.3: Photograph of exhaust gas heat exchanger gas side installation

1.2.2. Liquid side installation

There are fewer constraints on the liquid side of the heat exchanger both from a layout perspective and from the heat transfer medium. Oil and coolant were considered for use in the heat exchanger, however it was decided that coolant would be better because of the larger work required to pump oil, especially at lower temperatures when viscosity is high. Two circuit layouts were considered:

- An independent coolant circuit.
- Integrated into the main circuit.

With both designs, the aim was to promote engine warm-up and notably oil warm-up for reduced friction. To this end, the heat transferred from the exhaust gases to the coolant was directed to the engine oil cooler. In both cases the base circuit is the same as described in chapter 4, and assessed in detail in chapters 5 and 7.

(a) Secondary circuit

Figure 8.4 shows the layout of the coolant circuit with the exhaust gas heat exchanger installed as an independent circuit. Coolant was pumped using a variable speed electric pump between the exhaust gas heat exchanger and the oil cooler. For simplicity, the electric pump was driven from an independent power source as its power consumption would be small.

A header tank (not shown) provided sufficient pressure at the pump inlet. The header tank was essentially another degas bottle from the same engine and a significant volume of coolant was stored in this bottle. In total 4L of coolant was used in the secondary circuit, although only half of this volume was active during warm-up.

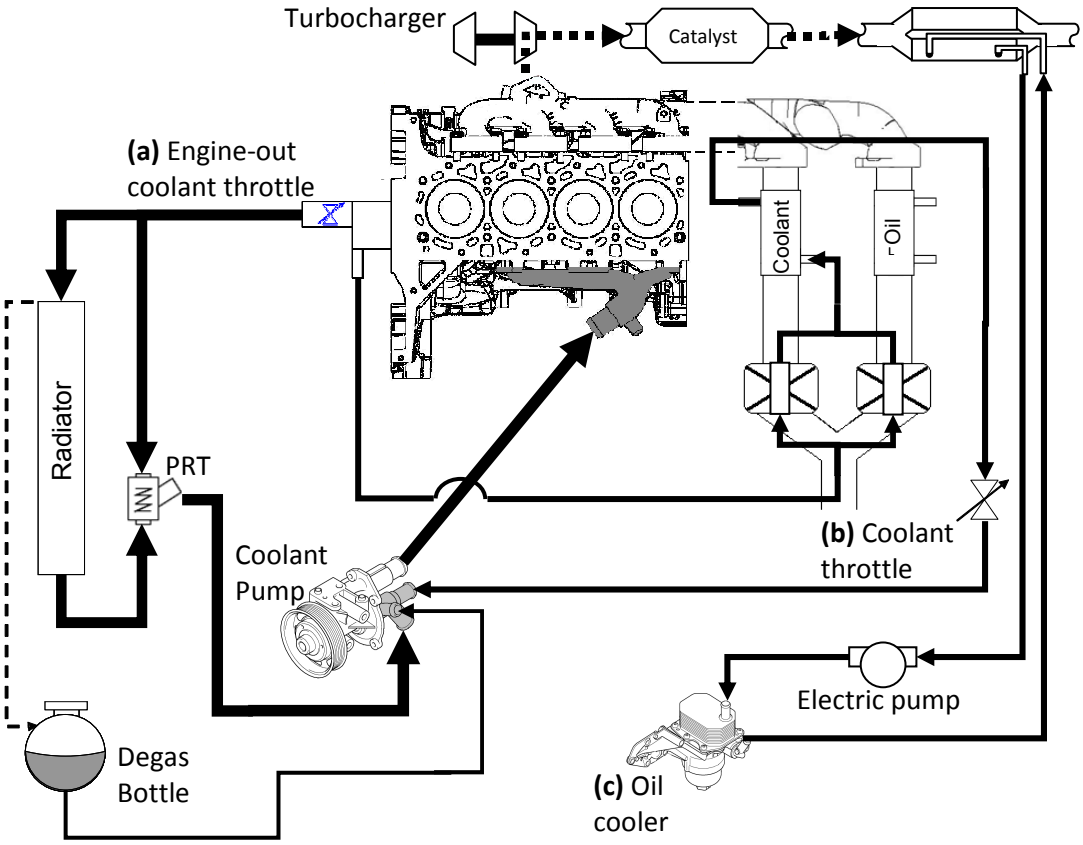


Figure 8.4: Coolant circuit layout for phase 1 with exhaust gas heat exchanger fitted into independent cooling circuit

The secondary circuit was instrumented with an ultrasonic flow meter and thermocouples to measure temperature differences over each of the heat exchangers. A thermocouple was also installed at the exit of the header tank to show any rises in temperature in that part of the circuit.

(b) Integrated into primary circuit

Figure 8.5 shows the coolant system layout with the exhaust gas heat exchanger integrated into the main cooling circuit. The heat exchanger is fitted into the EGR cooler loop such that coolant flowing in that loop will pick up heat from the EGR gases, then the exhaust gases before flowing into the oil cooler. The aim of this installation is to use the heat exchanger as an additional heat input to the coolant just before entering the oil/coolant heat exchanger. Coolant flow through the heat exchanger, coolant-cooled EGR cooler and oil cooler are controlled by a single throttle and can bypass the oil cooler using a 4 way valve. Because the circuit was integrated into the main circuit, there was no need for an additional pump and header tank and the extra coolant volume was only that in the two hoses to and from the heat exchanger. The total additional coolant volume was about 2L. Again the circuit was instrumented with thermocouples measuring temperature changes over the heat exchangers and ultrasonic flow meters measuring coolant flows in each leg.

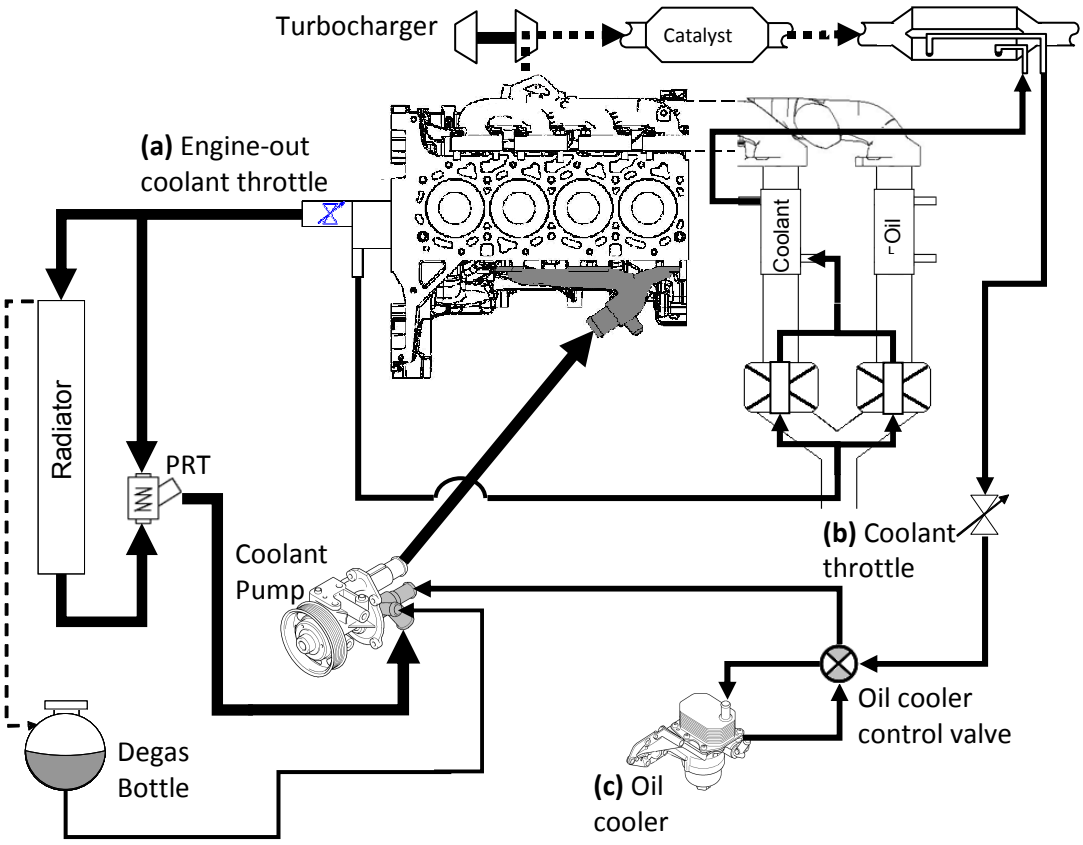


Figure 8.5: Coolant circuit layout for phase 2 with exhaust gas heat exchanger integrated into the main cooling circuit

2. *Results and system behaviour*

2.1. Heat exchanger behaviour

The heat exchanger performed similarly in both circuits. The heat exchanger was effective at extracting heat from the exhaust gases as shown by the gas temperature drop over the device shown in figure 8.6.

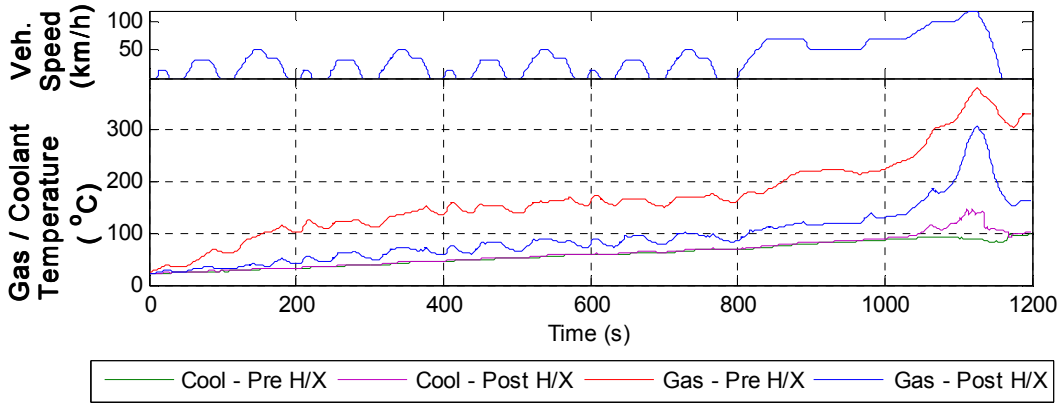


Figure 8.6: Gas and coolant temperatures at inlet and outlet of exhaust gas heat exchanger

The measured gas and coolant temperatures can be used to calculate the effectiveness of the heat exchanger detailed in equation 8.1 - [154]. The effectiveness *ε,* is the ratio of the actual heat transfer to the maximum possible heat transfer if the exhaust gas temperature leaving the heat exchanger could be reduced to that of the coolant entering the heat exchanger[85].

$$\varepsilon = \frac{q}{q_{\max}} = \frac{C_h\left(T_{h,i} - T_{h,o}\right)}{C_{\min}\left(T_{h,i} - T_{c,i}\right)} \qquad \textbf{8.1}$$

Where

- *C_h is the heat capacity of the hot gas (exhaust gases) calculated from the exhaust mass flow rate and exhaust gas specific heat capacity*
- *C_{min} is the smallest heat capacity of the two fluids (exhaust gas and coolant); in this case it is always the exhaust gas*
- *$T_{h,i}$ and $T_{h,o}$ are the hot fluid (exhaust gas) inlet and outlet temperatures respectively and $T_{c,i}$ is the cold fluid (coolant) inlet temperature.*

[85] In principle, such a heat exchange could be achieved in a counter flow device of infinite length [154].

To understand the relationship it is important to highlight the concept of minimum and maximum fluid. The minimum fluid is the fluid with the lowest heat capacity that will undergo the largest temperature change in the heat exchanger; it is effectively this fluid that limits the heat transfer from an energy perspective and is used as the denominator in equation 8.1. If thermal losses are neglected, the actual heat transfer can be calculated either on the gas or the coolant side. In practice, and notably from cold start, the warm-up of the heat exchanger components can represent a significant proportion of the heat, and in this case the effectiveness calculated on the gas side will be significantly higher than that calculated on the coolant side. The calculated heat exchanger effectiveness using both the gas and coolant side temperature drops are shown in figure 8.7.

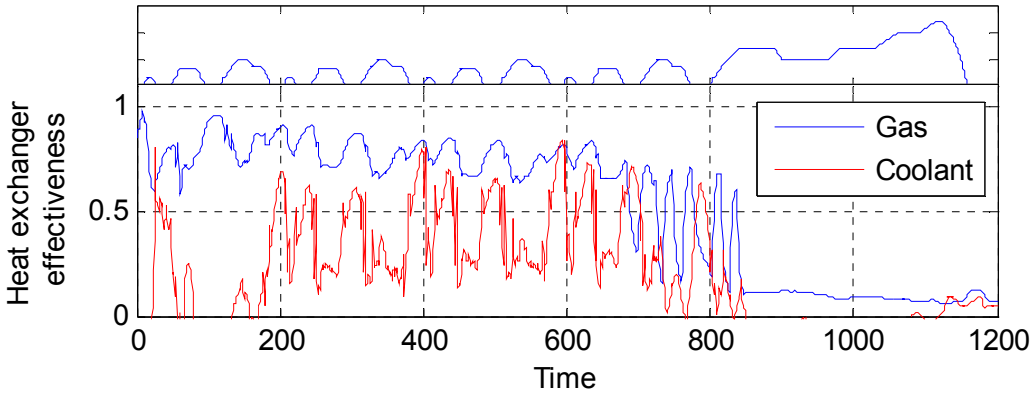


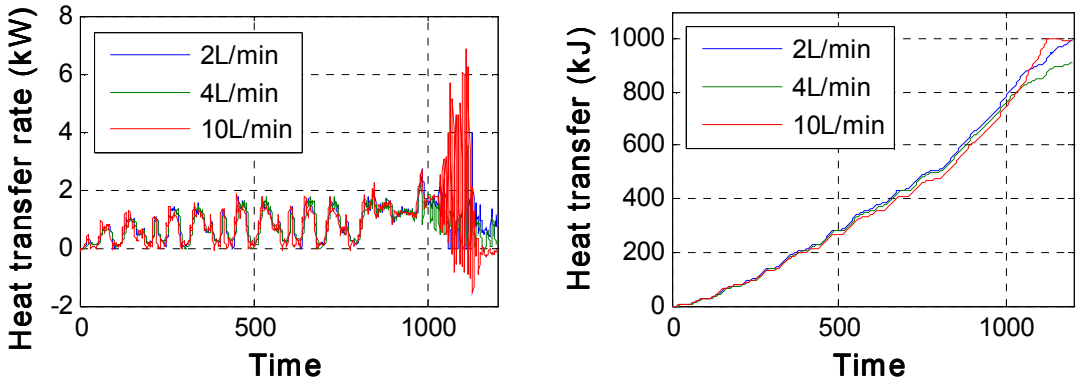
Figure 8.7: Exhaust gas heat exchanger effectiveness over NEDC based on measured gas and coolant temperatures and flow rates

- Based on gas side measurements, over the first 700 seconds of the drive cycle the effectiveness remains high, varying between 0.65 and 0.95 with a tendency to drop as the drive cycle progresses. This can be explained by the increase in metal temperature of the heat exchanger which reduces the thermal gradient between the gases and the metal, reducing convective heat transfer. Over the following 150 seconds, there is a transition period as the coolant temperature fluctuates around 90^{o}C and the simple on/off control algorithm successively engages and bypasses the heat exchanger on the gas side. After 850 seconds the bypass valve isolates the heat exchanger and heat transfer is inhibited. It is important to note that the effectiveness calculations are meaningless when the valve is bypassed as although exhaust gases are flowing through the heat exchanger device, they are not flowing over the heat exchanging surface.

- The effectiveness calculated on the coolant side is significantly lower. In addition, over the first 150 seconds, the values are essentially meaningless as the temperature difference between inlet and outlet is so small. Although the calculated effectiveness is low, the overall trend is opposite to that seen on the gas side with a tendency to increase as the components warm-up. The hotter metal increases the thermal gradient between heat exchanger and the coolant, favouring convective heat transfer.

Because the aim of the heat exchanger is to warm up the coolant rather than cool down the gas, it would seem sensible to refer to the effectiveness calculated using the coolant side data. However the small temperature differences make this difficult as the relative temperature difference is uncomfortably close to the measurement accuracy of the thermocouples. Later in the drive cycle, the effectiveness calculated in this way is significantly lower than that from the gas side calculation with the difference incorporating the heat losses but also the warm-up of the heat exchanger itself.

Figure 8.8 shows a measure of the available heat extracted from the exhaust gas for different coolant flow rates. This clearly shows the fluctuations in available exhaust heat depending on the engine operating condition. There is also a small warm-up of the exhaust pipe components over the first 200 seconds. Approximately 500kJ can be collected over phase 1 and up to 1MJ over the NEDC.



(a) *(b)*

Figure 8.8: (a) Heat transfer rate and (b) heat transfer in exhaust gas heat exchanger and oil cooler

The variation in coolant flow rate had little effect on heat transfer. From a theoretical perspective this would be expected as the heat transfer coefficient in the heat exchanger would be dominated by the gas side convection, meaning changes to the coolant flow would

cause little difference compared to changes in exhaust gas flow. This is clearly the case in figure 8.8 (a) where there is little difference between the curves, but significant changes due to engine operating condition.

2.2. Systems level behaviour

Initially the secondary circuit was installed and despite the significant amount of heat available from the exhaust gases, the coolant temperature in the secondary circuit remained colder than the oil temperature. This can be seen in figure 8.9 where the coolant temperature in the secondary circuit at the inlet to the oil cooler is lower than the oil temperature entering the oil cooler for the majority of the warm-up period. The second law of thermodynamics dictates that heat transfer from coolant to oil is impossible, and in fact heat transfer will occur in the opposite direction, effectively cooling the oil during warm-up. In practice the temperature difference is quite small and the resulting heat transfer will also be small, but clearly this is not the intended behaviour of the system.

It is important to note that even though there is no heat transfer from coolant to oil, the circuit may still be beneficial as the heat from the coolant will contribute to the warm-up of the oil heat exchanger metal, which would otherwise require heat from the oil. However the results remain disappointing.

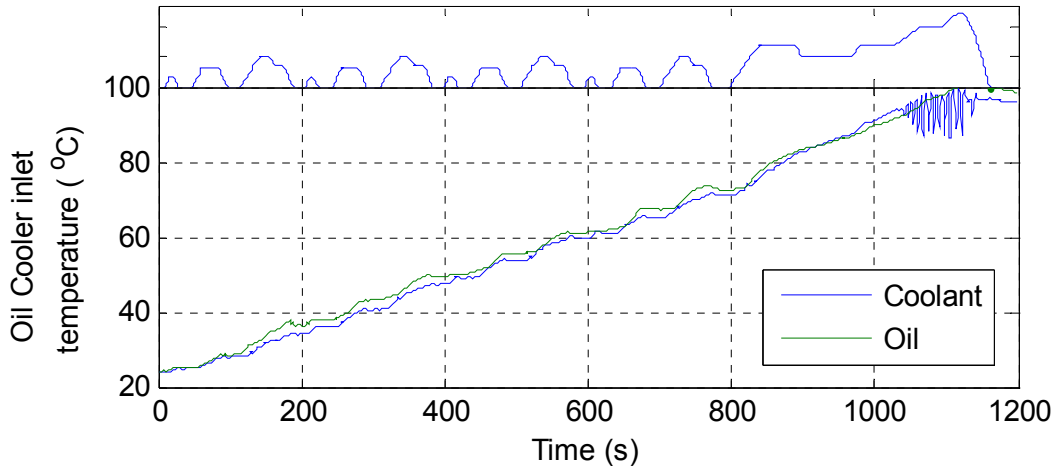


Figure 8.9: Coolant and oil temperatures at inlet to oil cooler for system using exhaust gas heat exchanger in secondary circuit

This behaviour is a result of the energy available from the exhaust gases being slightly less than the energy required to heat the additional hardware. Two simple simulations were performed to show the limiting effects of the added thermal inertia:

1. Using the measured heat input from the exhaust gas heat exchanger, the temperature of different masses of water was calculated assuming no other heat losses. The results are compared to the measured oil temperature during warm-up in figure 8.10 (a).
2. Building on the previous simulation, heat transfer to oil and heat loss to ambient air were simulated using empirical values for heat transfer coefficients. This simulation does not include calculation of oil temperature, but gives a more realistic water warm-up profile. The thermal inertia of the system is again treated as a mass of water and results are compared to a measured oil temperature in figure 8.10 (b).

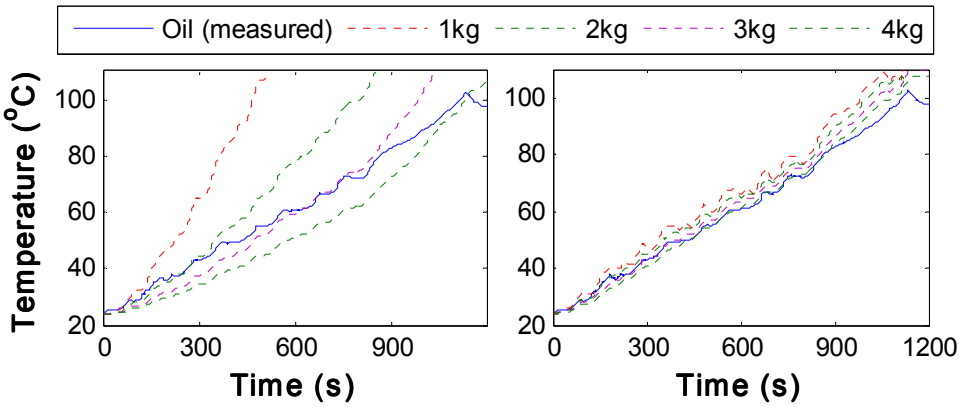


Figure 8.10: Simulated water temperature for different water masses using heat input from exhaust gas heat exchanger with (a) assuming no heat losses and (b) assuming simulated heat transfer to oil and ambient air. Results are compared to measured oil temperature during warm-up.

The simulations clearly show that to achieve heat transfer to oil the system inertia must be low. This is difficult because of on-vehicle constraints for exhaust after-treatment which push the heat exchanger further down the exhaust pipe[86]. To achieve any significant benefits it was estimated that the additional inertia should be below that of 2kg of water. The requirements on inertia are even more demanding to achieve positive heat transfer to oil over the first 500 seconds. The results from presented in figure 8.9 are explained by an approximated thermal inertia equivalent to 3.3kg of water when including the pipework and components[87]. In the

[86] Components like particulate filters and catalysts need to be installed close to the engine as they require combustion heat for regeneration and light off.

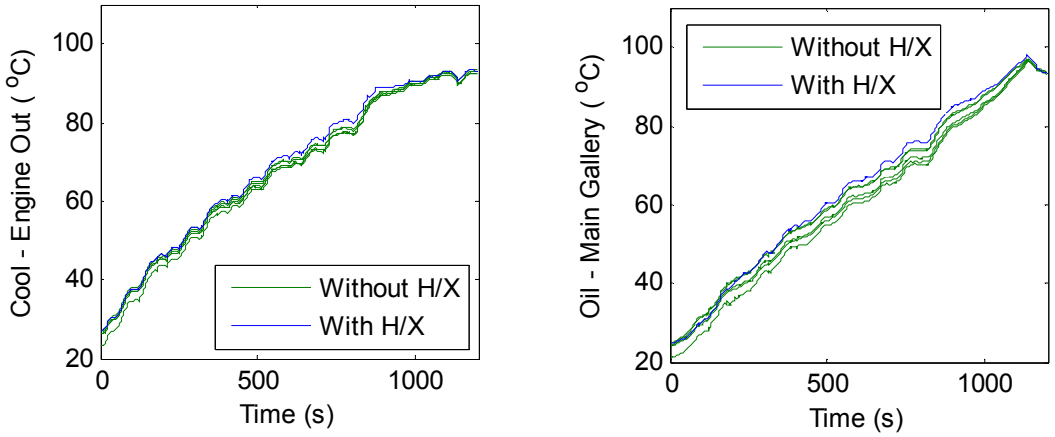
[87] The additional components included the heat exchanger itself, the electric pump and an ultrasonic flow meter for instrumentation purposes.

system implemented by Andrews *et al.* [46], the exhaust gas heat exchanger was installed between the exhaust ports and exhaust manifold. This configuration does not account for after-treatment devices, but give maximum exhaust heat with virtually no increase in coolant circuit inertia. Clearly this explains the difference in performance with the results from this work.

Thermal inertia was reduced slightly with the integrated circuit as the electric pump and flow meter could be removed along with a small reduction in hose length. Each of the active TMS devices remained available and the effect of the heat exchanger was tested with and without engine-out coolant throttle mapping[88]. These calibrations are described in table 8.1 and coolant and oil temperatures are shown in figures 8.11 and 8.12 for open and mapped engine-out coolant throttle respectively. In figure 8.11 the multiple lines represent repeat tests at the same condition, however these were not available at the condition in figure 8.12.

	1	**2**
Engine-out coolant throttle	Open	Mapped
EGR cooler coolant throttle	Max Flow	
EGR cooler type	Coolant	
Oil cooler coolant flow	On	

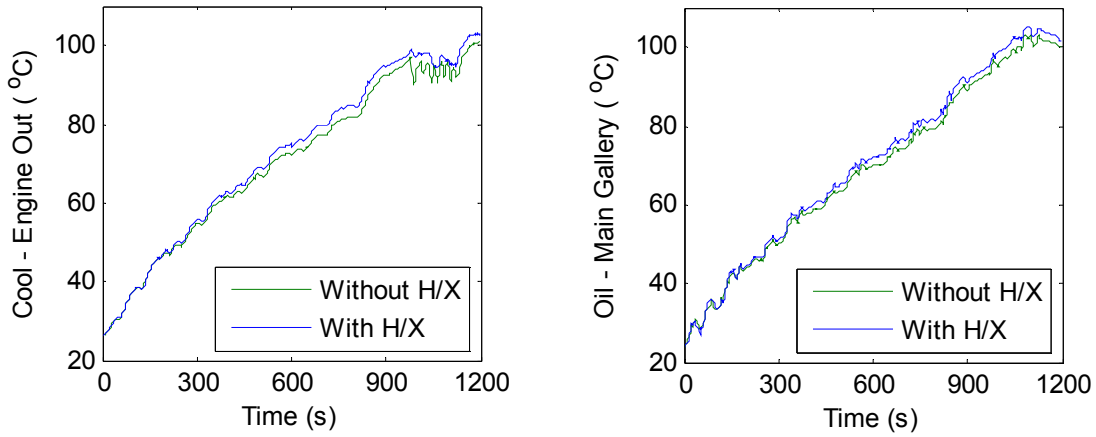
Table 8.1: Two setups for comparison with and without exhaust gas heat exchanger



(a) *(b)*

Figure 8.11: (a) Coolant and (b) oil temperatures for engine with and without heat exchanger hardware with engine-out coolant throttle open

[88] The tests without the exhaust gas heat exchanger refer to a configuration without all of the additional hardware, not a case where the hardware is installed but heat exchanger bypassed. These results were issued from the DoE experiments described in chapter 7.



(a) *(b)*

Figure 8.12: (a) Coolant and (b) oil temperatures for engine with and without heat exchanger hardware with engine-out coolant throttle mapped

In both cases it can be seen that the coolant and oil warm-up rate were faster than with the exhaust gas heat exchanger. The benefit is small and only apparent after the first 300 seconds of the drive cycle. Analysis of the coolant temperature and oil temperature in the oil cooler show that for this system there will be heat transfer between the coolant and oil (see figure 8.13)[89]. In previous chapters it has been shown that using the oil cooler during warm-up does provide heat transfer to oil and improve oil warm-up rate; the inclusion of the exhaust gas heat exchanger appears to improve this slightly.

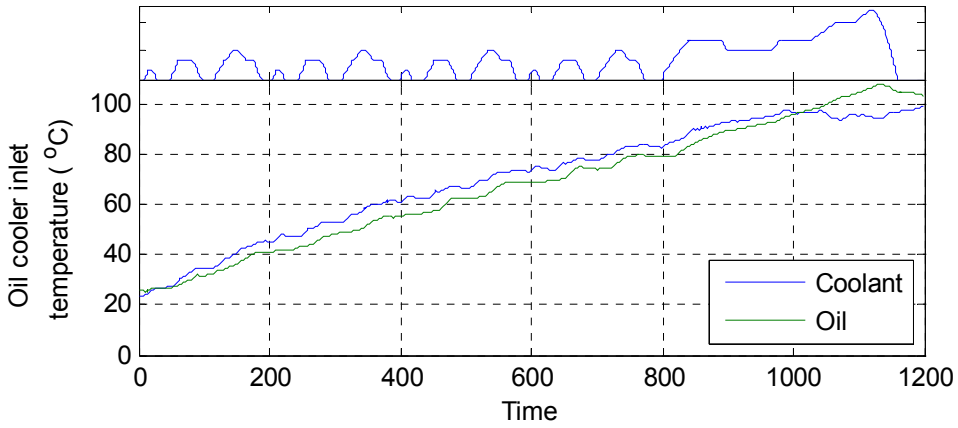


Figure 8.13: Coolant and oil temperatures at inlet to oil cooler for system using exhaust gas heat exchanger integrated into main coolant circuit

[89] There is significantly less noise in the coolant temperature when the exhaust gas heat exchanger bypass valve is use towards the end of the test compared to the installation in the secondary circuit: this can be explained by the larger mass of coolant in the main circuit that damps these fluctuations.

2.3. Detailed system analysis

2.3.1. Approach and test plan

As there appears to be a benefit, albeit small, from including the heat exchanger, it was decided to use a similar DoE approach to that in chapter 7. This would allow the interactions between in the system to be described and if necessary optimise the use of the hardware. Previous work had established that restricting the coolant flow in the engine was beneficial to warm-up throughout the engine so in this experiment the throttle would be controlled in all cases. Similarly, although the heat exchanger was fitted with a bypass valve, it would not make sense to bypass the heat exchanger if the additional coolant circuit is installed. In all cases the bypass valve was set to flow exhaust gases through the heat exchanger until coolant reached 90^{o}C. Because the changes in warm-up rate were likely to be small, SOI was not included in this experiment. The oil cooler, dual EGR cooler and second coolant throttle were all used as they have been proven to modulate heat flows between the various fluids. These actuators would allow the engine to operate in one of the following calibrations:

1. " All heat to oil" – using oil-cooled EGR and oil cooler.
2. "All heat to coolant" – using coolant-cooled EGR and bypassing the oil cooler.
3. A mix of the previous two setups.

Each test point on the experimental design would be a cold-start NEDC with the calibration fixed throughout the cycle. The three factors and factor levels were:

1. Coolant flow in the EGR/exhaust gas heat exchanger and oil cooler loop (flow settings at idle were 6L/min, 3L/min and 0.8L/min)
2. Oil cooler bypass valve (on or off)
3. EGR cooler type (coolant or oil)

Because only three factors were of interest, a full factorial design was used and the experimental test points are listed in table 8.2. As with the previous approach and because the changes in test points were only software changes, the test plan was run in a randomised order.

#	1	2	3
DoE Number	EGR/ H/X / Oil cooler coolant flow at idle (L/min)	Oil cooler	EGR valve
1	0.8	On	Coolant
2	3	On	Coolant
3	6	On	Coolant
4	0.8	On	Oil
5	3	On	Oil
6	6	On	Oil
7	0.8	Off	Coolant
8	3	Off	Coolant
9	6	Off	Coolant
10	0.8	Off	Oil
11	3	Off	Oil
12	6	Off	Oil

Table 8.2: Experimental design test points for assessing active thermal management system with exhaust gas heat exchanger

2.3.2. Modelling results

The same modelling approach was taken as described in chapter 7 with measured outputs modelled using simple polynomials. EGR type and oil cooler bypass valve were qualitative variables whereas the coolant flow could be varied continuously. As with the previous modelling work, the coolant flow was modelled against the volume of coolant in litres that flowed through the EGR/Heat exchanger/oil cooler loop over the appropriate period of the drive cycle. The analysis will focus on phase 1 of the NEDC as this is the phase where benefits from the heat exchanger are expected. For large parts of phase 2 the heat exchanger is bypassed as the coolant has reached its warm operating temperature.

Suitable models were fitted to the data to represent the temperature rises in different parts of the engine. The rise in mid stroke liner temperature and oil gallery over phase 1 are shown in figures 8.14 and 8.15 respectively. As was the case in the previous chapter, the system actuators offer a trade-off between upper and lower engine warm-up. This is of no surprise as the addition of the exhaust gas heat exchanger does not affect the behaviour of the other actuators, but simply allows additional heat to be put into the system.

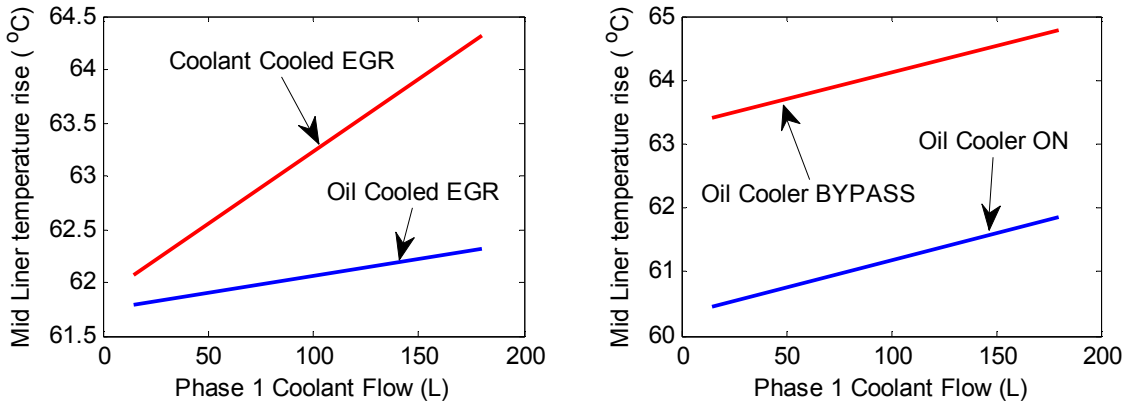


Figure 8.14: Effect of coolant flow, oil cooler and EGR cooler on mid stroke cylinder liner temperature rise over phase 1 of NEDC

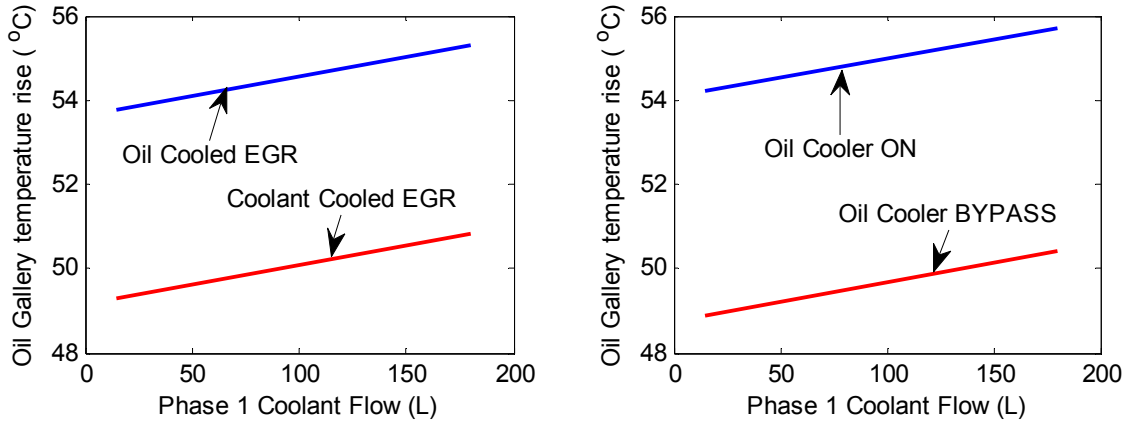
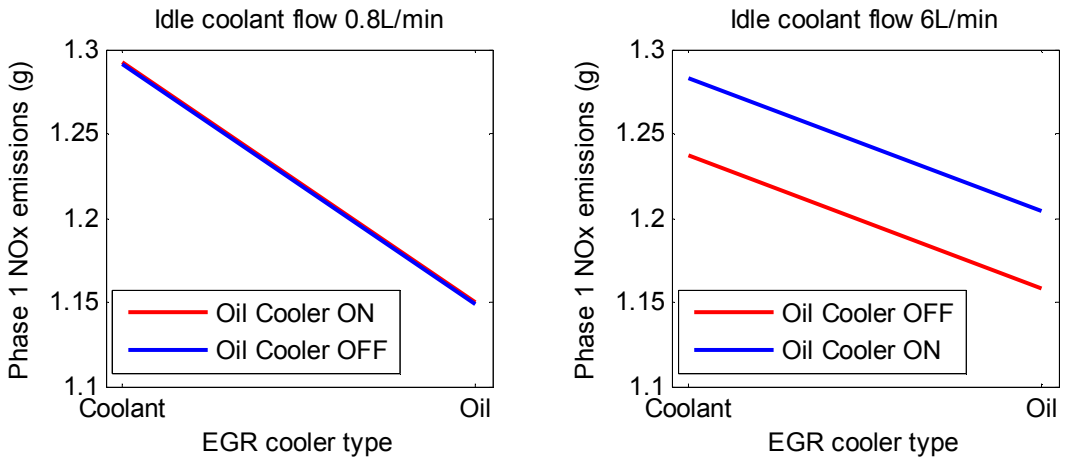


Figure 8.15: Effect of coolant flow, oil cooler and EGR cooler on oil gallery temperature rise over phase 1 of NEDC

It was attempted to model fuel consumption and emissions data. For the case of fuel consumption, no good-fitting models were found based on the input variables used in this experiment. When assessing the raw data no concrete trends were identified, suggesting that any variations in fuel consumption over this experiment are smaller than the random variations incurred in the testing.

Better fitting models were developed for CO and HC and NO_x emissions, although all of the variation in results was attributed to the switching of EGR gas cooler type. Using the oil-cooled EGR cooler increased HC and CO emissions but reduced NO_x. These effects were discussed in chapter 7 and are explained by changes the effectiveness of EGR gas cooling. In these experiments, some other effects were apparent on NO_x emissions as shown in figure 8.16. Both the coolant flow rate and the oil cooler affected NO_x emissions. The lowest NO_x were achieved by using oil-cooled EGR in combination with coldest possible oil. The colder oil was achieved by reducing or inhibiting coolant flow in the oil cooler during warm-up.



(a) *(b)*

Figure 8.16: NO_x variation for oil cooler and EGR cooler type calibration at (a) low coolant flow and (b) high coolant flow

3. *Discussion and overall system benefits*

The heat exchanger was not expected to significantly alter the behaviour of the active components during warm-up and as a result the opposing trends in upper and lower engine were not surprising. However, the inclusion of the exhaust gas heat exchanger is aimed at pushing this trade-off beyond the Pareto front measured in the previous chapter. Figure 8.17 shows the familiar warm-up trade-off resulting from the active thermal management (ATM). It was shown that throttling engine-out coolant flow could improve warm-up throughout, however the oil cooler and EGR controls produced a trade-off between upper and lower engine temperatures. The results from this chapter have now been superimposed on the same axes. The different oil cooler and EGR cooler calibrations are grouped and the effect of coolant flow reduction is illustrated by the blue arrows.

It is clear from figure 8.17 that using the exhaust gas heat exchanger has improved upon this trade-off. A new Pareto front has become apparent with approximately 1-2^{o}C benefit in both upper and lower warm-up over phase 1. It is also interesting to note that the exhaust gas heat exchanger has extended the range over which the trade-off may be varied, notably in favour of oil warm-up. With the previous system the largest oil temperature rise over phase 1 was around 53^{o}C, however with the heat exchanger 57^{o}C is achieved.

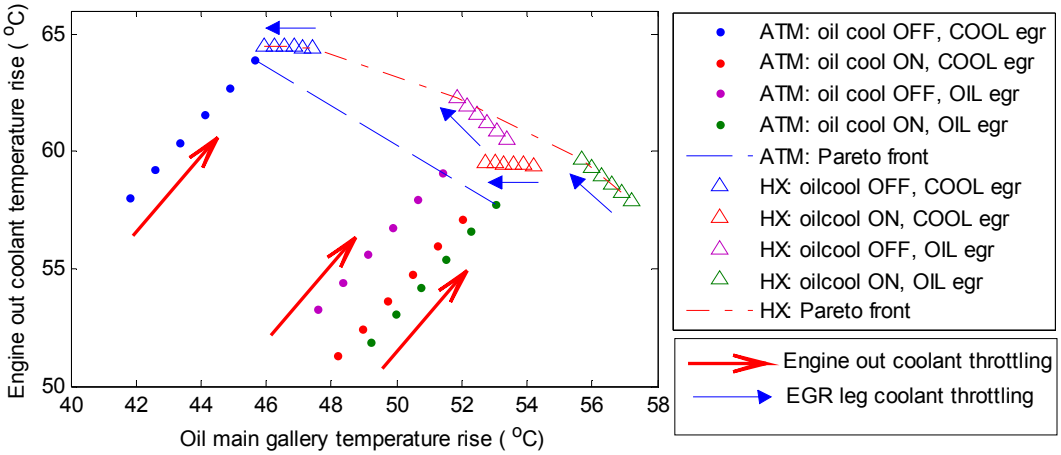


Figure 8.17: Coolant and oil warm-up trade-off for active thermal management (ATM) and heat exchanger (HX) hardware builds

The response model for NO_x suggested that varying coolant flow rates in the EGR cooler/oil cooler loop also affected these emissions. It is interesting that this effect was not apparent in the results from the DoE in the previous chapter. This could be explained by two factors: On the one hand this could be a result of the exhaust gas heat exchanger amplifying the effect because of the additional heat available. Alternatively, this could simple be apparent because of the smaller variations and the larger number of test points compared to the number of inputs of the experimental design[90].

The spread of carbon balance fuel consumption and NO_x emissions for the two DoE campaigns on the active thermal management[91] (ATM) and exhaust gas heat exchanger (HX) are shown in figures 8.18 and 8.19 respectively. In each case the results are presented as box plots showing the median and data spread; results are also broken down into phase 1 and NEDC. Although these plots are commonly used to display repeatability in this case they are showing the intentional data spread from the DoE. The spread is larger for the ATM results because the perturbations of the input variables were larger.

[90] In this chapter a full factorial design was used with 12 test points for determining the effects and interactions of 3 input variables; in the previous chapter 17 test points were used to determine the effect and interactions of 5 inputs. An analogy could be made to the resolution of fractional factorial designs [112] which is dependent on the smallest interaction order that is confounded with a main effect.

[91] This refers to the configuration optimised in chapter 7.

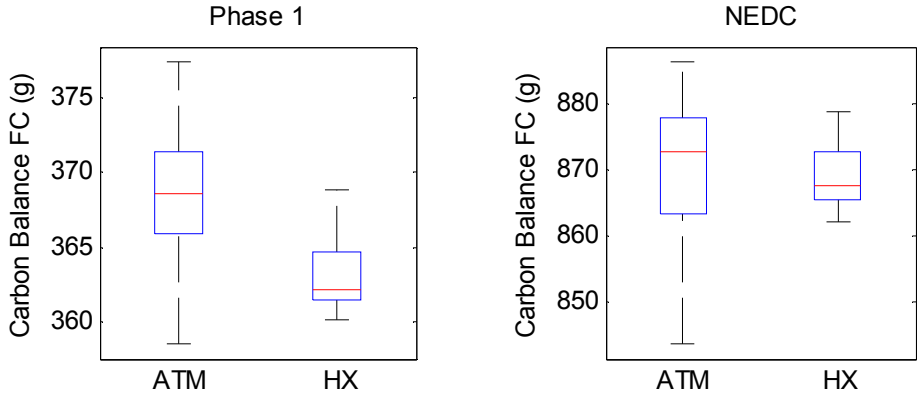


Figure 8.18: Carbon balance fuel consumption for all DoE test points for both active thermal management (ATM) and heat exchanger (HX) hardware configurations (a) phase 1 and (b) NEDC

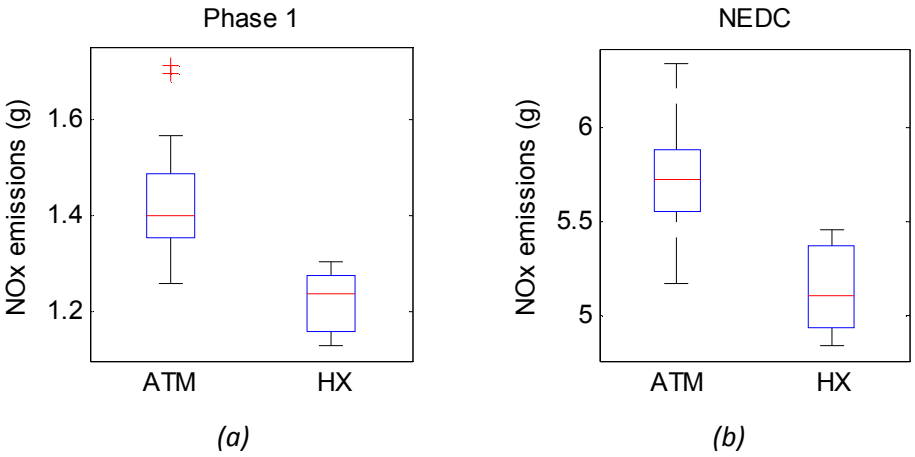


Figure 8.19: NO_x emissions for all DoE test points for both active thermal management (ATM) and heat exchanger (HX) hardware configurations (a) phase 1 and (b) NEDC

It is clear from these results that there is no significant FC benefit from the exhaust gas heat exchanger. For phase 1 the fuel consumption from the heat exchanger setups sits towards the lower end of the ATM tests, which is explained by the closed engine-out coolant throttle. When phase 2 is considered, FC results with the heat exchanger lie in the upper part of the ATM results because of differences in injection timing. There is a clear step change in measured NO_x emissions, however if the empirical NO_x correction factor derived in chapter 3 is applied, the difference in NO_x is almost entirely removed. Although there appeared to be small benefits in terms of thermal behaviour with the exhaust gas heat exchanger, these have not been reflected in FC and NO_x. Based on the simulated thermal behaviour previously described in this chapter, significant design work would be required to achieve large enough variations in thermal to be reflected in engine performance.

4. *Chapter summary and conclusions*

In this chapter an exhaust gas heat exchanger was installed into the thermal management system. The aim was to use otherwise waste heat to improve engine warm-up, however the impact was not as significant as expected. A DoE approach similar to that presented in chapter 7 was conducted and results were compared to those from the system without the heat exchanger. The following conclusions can be drawn from this exercise:

1) The heat exchanger itself had a high effectiveness, although heat transfer was dominated by changes in exhaust flow. This means the benefit of the device is strongly dependent on duty cycle. Over the NEDC, 1MJ of heat was extracted, although a significant proportion of this is only available in phase 2 when warm-up is less critical.

2) In practice, the main limitation stems from requirements of exhaust gas after-treatment devices. With current and foreseeable legislation and technology development trends, it is difficult to see an exhaust gas heat exchanger taking priority over these devices. This impacts both the available exhaust gas temperature and the increased thermal inertia required to integrate the device in the cooling circuit.

3) It is the large additional thermal inertia that compromises the system and the available exhaust heat was only just enough to overcome this inertia. To give significant benefits, the equivalent additional inertia should be less than that of 1kg of water.

4) The trade-off between upper and lower engine was again apparent due to changes in oil cooler and EGR calibration. A small benefit in the Pareto front of 1-2^{o}C was measured but this did not have any significant effects on fuel consumption and emissions.

5) Because of the strong dependency on duty cycle and the focus of this work on engine warm-up, these tests may not make best use of the heat exchanger. In real world, full-vehicle applications, cabin heating, other powertrain element and lower start temperatures may show a more viable application.

Chapter 9 - *Global engine modelling during warm-up*

The previous chapters have demonstrated the benefits and performance of new engine hardware for improved under warm-up. Together they have demonstrated a method for optimising the use of these devices over a specific duty cycle, when the engine hardware is available. However there are many applications during engine development when hardware is not available and engineers must rely on simulation results. In the literature review, lumped thermal capacity models were described which, when correctly configured, allow the simulation of engine temperatures during thermal transients. In this final working chapter, a global, dynamic, data-driven model will be developed to use in conjunction with these lumped capacity models.

Dynamic models are emerging as an expansion of steady state models because they allow characterisation of the engine under dynamic conditions with much reduced data acquisition times compared to conventional steady state approaches. The technique is still at infancy and has so far failed to capture the thermal behaviour of the engine. This chapter will produce dynamic models for NO_x and CO_2 emissions and build upon the work of previous studies to include thermal aspects.

The work presented in this chapter was presented at the 6^{th} conference of Design of Experiments in engine development conference in Berlin, May 2011 [155].

1. *Global and thermal modelling approach*

The concept of dynamic design of experiments was introduced in chapter 3 and this methodology was applied to the PUMA 2.0L engine because it lent itself better to the control of parameters than the PUMA 2.4L engine. Although this is a significant compromise in terms of the available instrumentation, it did allow the methodology to be demonstrated within the time frame of this thesis. The aim was to capture and model dynamic characteristics of NO_x emissions and fuel consumption over the operating region of the NEDC

The first aspect of this chapter aims at reproducing the work of other authors [119] to capture and model the dynamic behaviour of the engine. Although in itself is not novel, this work will take a more practical perspective compared to the more modelling side perspectives seen in the literature. A selection of calibration variables will be chosen for this aspect in line with other published work and Volterra series models were to be fitted. The test plan was designed around the NEDC in terms of range and the validation sequence was NEDC.

The main aim of this chapter is to build on this dynamic modelling approach to incorporate thermal aspects of the engine behaviour. As mentioned in chapter 3, unlike common calibration variables, the engine temperature cannot be varied freely according to a prescribed input signal because it is not possible to reproduce a realistic warm-up temperature distribution. A simple approach to include the thermal behaviour is through the inclusion of a temperature scaling factor applied to the Volterra model. This modelling approach is presented graphically in figure 9.1 and expressed mathematically in equation 9.1.

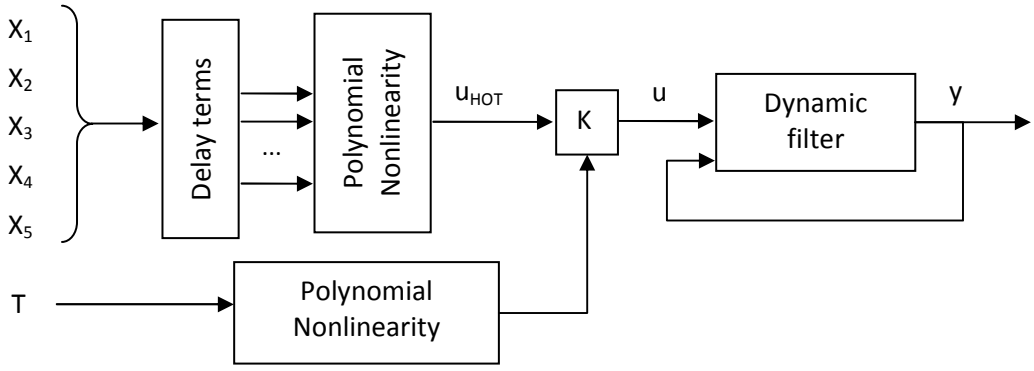


Figure 9.1: Graphical overview of Volterra series model with temperature based scaling factor

$$\hat{y}(k) = f(T) \times \begin{pmatrix} f([x_1(k), x_2(k), \ldots, x_n(k)]) \\ + f([x_1(k-1), x_2(k-1), \ldots,], [x_1(k-2), x_2(k-2), \ldots], \ldots) \\ + f(\hat{y}(k-1)) \end{pmatrix} \quad \textbf{9.1}$$

All model input excluding temperature are fed into the main Volterra model. Some delayed input terms can be fed forward to account for the system dynamics. A global polynomial equation of up to 4th order and including interaction terms account for the non-linearity. This is scaled by a factor dependant on temperature input before an autoregressive aspect accounts for sensor dynamics.

The models produced in this work will predict specific ouputs, i.e. emissions measured in ppm or % and fuel consumption in g/kWh. This approach is taken because the actual fuel consumption and emissions correlate very strongly with engine speed a torque and the effect of calibration variables is hard to extract. By modelling specific quantities the direct effects of speed a load are removed and the model is aimed at giving information on the efficiency of the engine at any particular operating point.

2. *Engine control and test plan*

2.1. Engine control

2.1.1. Input signals and control

(a) Engine speed and load

Engine speed and load were varied over the desired range while a number of other ECU calibration variables were also excited. The engine speed was directly controlled by the dynamometer; the engine load was controlled by electronic pedal actuation that could either be part of a feedback system with a torque set-point or a direct pedal set-point. It is important to compare the expected behaviour these two models will have with respect to changes in thermal state of the engine over the validation NEDC:

- The torque based model will have the same inputs in terms of load irrespective of thermal state because to input refers to the brake output of the engine;

- Conversely, the pedal based model will have different inputs depending on the engine thermal state. This is because the model input relating to engine load (i.e. pedal position) is more closely assimilated with the quantity of fuel injected and hence the indicated work. This will be higher under colder conditions due to increased frictional losses

These aspects are important when considering the role of the temperature scaling factor:

- for the torque based model, the scaling factor needs to account for changes in thermal state **and** changes in fuelling
- for the pedal based approach the scaling factor only needs to account for changes in thermal state.

As a result, it was decided that both the torque and the direct pedal approaches would be tested independently for comparison.

(b) ECU control variables

Initially four other ECU variables were considered as inputs for the modelling exercise[92] and all other variables were left to the production calibration: these four variables are listed in table 9.1 with their excitation methods. It was necessary to investigate the ECU architecture to find areas where independent control of these values could be taken using the calibration interface. Within the ECU architecture, some variables allow direct setting of parameters whereas other variables allow for easy control relative to a base calibration, either by 'adder' or 'multiplier' functions. As the names suggest, an adder function allows an offset to be added manually to the ECU generated demand for a specific actuator whilst a multiplier allows the ECU demand to be scaled by a manually-set factor. Each of these capabilities has different applications during the calibration process, or they can be combined by multiplying by zero and then offsetting to any desired value. It is fortuitous that these features can be addressed rapidly by an appropriately configured host system to assume transient control over engine actuators without needing to resort to modified code or prototype hardware.

[92] A number of other variables could be chosen, notably describing the injection profile (pilot timing, quantity), however the number of variables was kept small as this was a first attempt to demonstrate the use dynamic modelling approach.

Variable	Excitation method
Injection timing	Direct acting 'adder' function; offset to production calibration was set directly in ^{o}CA
VGT position	Upper and lower limiter controls; set-point was controlled by setting upper and lower limit to desired set-point
Common Rail fuel pressure	Direct acting 'adder' function; offset to production calibration was set directly in bar
EGR valve position	Indirect control through Mass air flow set point using 'adder' function. This was necessary as the EGR valve position itself is controlled closed loop according to a MAF demand.

Table 9.1: Calibrations variables considered in this work and their excitation method within the ECU calibration

2.1.2. Control issues

The technical challenge of accurately controlling the engine according to the specified input signals and remaining within the safe operating limits of the hardware should not be underestimated. A number of issues were encountered during this work and are detailed in this section.

The low bandwidth of the ASAP3 connection compromised the study as using direct set-points was not fast enough to follow other dynamics in the engine. This affected the ability to vary the VGT position because the upper and lower limiter values could not be updated fast enough to follow the changes in engine speed and load. This is illustrated in figure 9.2: the left frame shows the VGT position for a given speed and load excitation. In blue is the VGT position demand that the ECU would provide without any additional control and superimposed in red is the resultant demand signal when the host system is in control via ASAP3 link. The signal programmed into the host system is exactly the same as the previously measured ECU demand. It is clear that the signal is affected in two ways: firstly the signal is discretized because of the low 1Hz update frequency[93] of the ASAP3 link and secondly a lag is introduced. Physically this causes excessively high exhaust back-pressures as the VGT is forced to close when the production calibration would request an open setting. As no simple solution to this problem could be found, it was decided to reduce the scope of the experiment and not include the effect of VGT position. Whilst this was unfortunate, it did not affect the primary goal which aimed at modelling behaviour during warm-up.

[93] Within the host system the update frequency was significantly higher and the low update frequency was only caused by the ASAP3 connection.

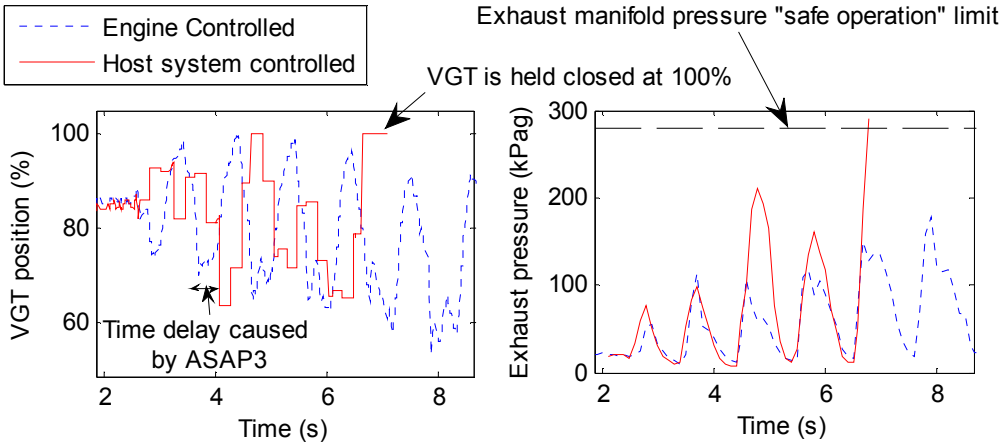


Figure 9.2: VGT actuation with and without control over ASAP 3 showing how the low update rate affects exhaust manifold pressure leading to unsafe engine operation.

The low update frequency of the ASAP3 link did not affect the other variables as they could be controlled through the "adder" and "multiplier" functions. This meant that the fast response of the ECU could be retained while the lower frequency input signals were applied downstream in the control algorithm. To prove the valid response of these variables to the adder functions, two repeat experiments were run: the first only exciting speed and load, the latter exciting in addition the appropriate control variable. After alignment of the data from each tests, the signals from each test were used to reconstruct the original input signal. In each case this gave good results as illustrated for rail pressure in figure 9.3

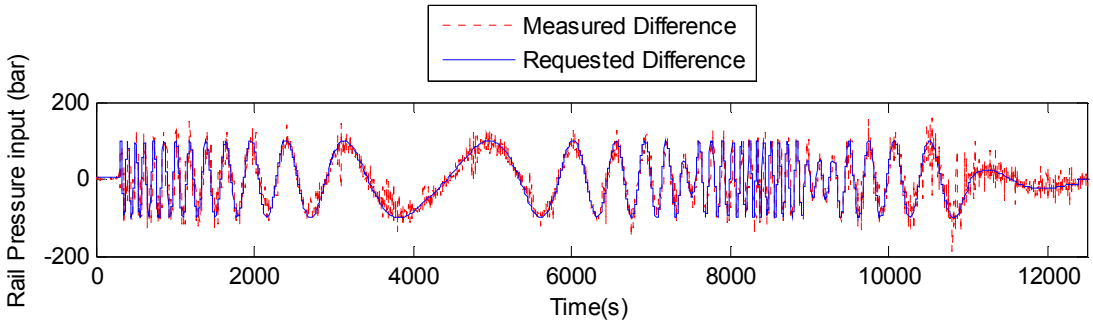


Figure 9.3: Input and measured excitation for rail pressure showing good agreement

An ideal input signal should cover the complete operating range in speed and load of the NEDC whilst giving a good range of the other calibration variables. As the NEDC includes idle phases, the requirement is to cover a speed range down to 800rpm and effectively 0Nm. The test control system encountered difficulties in covering these low speeds and torques with a single test because of the sensitivity of the pedal around the zero load point. The engine requires a

minimum pedal position to maintain a certain operating speed and below this the ECU idle controller takes over and ensures the engine does not stall. During scoping testing for running the chirp excitations, it was observed that at low torque demands[94], the engine would fail to follow the requested speed (see figure 9.4). To account for this problem, the experiment was split into two parts: A *normal running* phase and an *idle* phase.

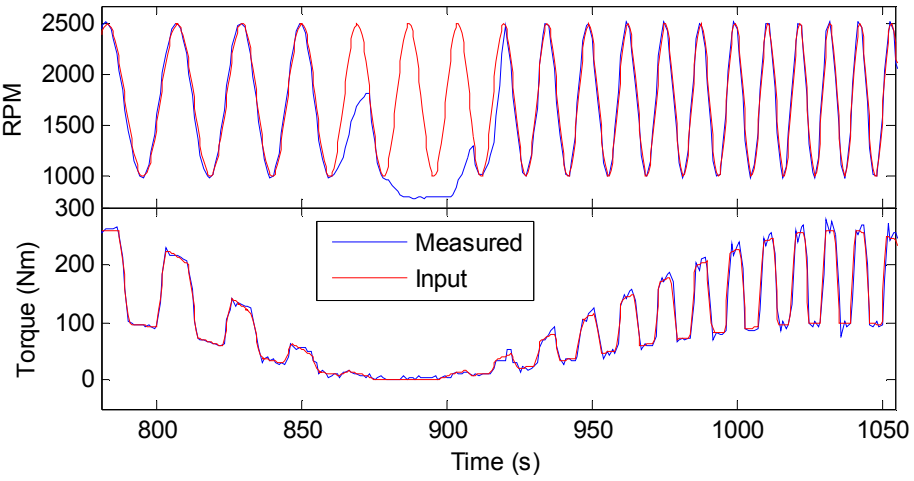


Figure 9.4: Engine speed and torque showing idle takeover at low torques.

2.2. Excitation signals

2.2.1. Capturing fully warm engine behaviour

For the main training data set, swept frequency sinusoidal input signals or chirps [156] were used[95]. The training data test sequence is shown for the torque based excitation in figure 9.5. The split between normal running and idle phases is clearly visible. The base chirp signals were produced using the integrated Mathworks Matlab '*Chirp*' function which creates a cosine function with varying frequency. The frequency was varied logarithmically as this best approximated the NEDC frequency range. These chirp signals, typically 10-20mins in length were assembled into an hour long sequence with the phasing of the inputs and the lower limit frequency optimised to give best design space coverage [119].

[94] For the pedal based approach this represented low pedal position demands.
[95] In chapter 3 different excitations signals were presented and the benefits and problems in the context of engine modelling were discussed. Chirp signals were concluded as being most useful for global engine modelling where speed and load would be used as inputs.

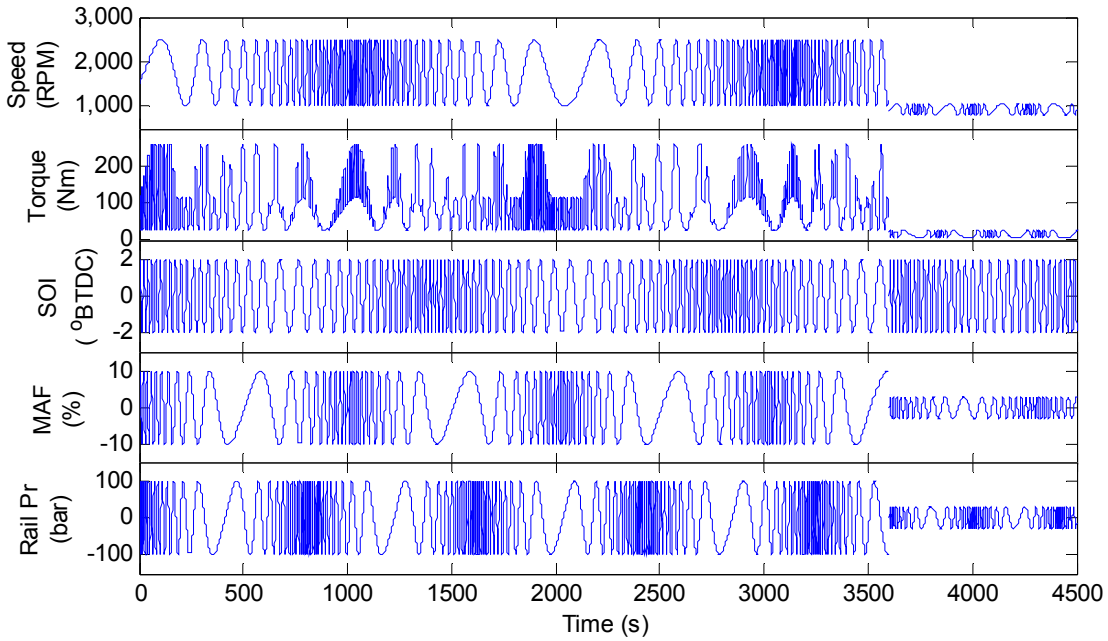


Figure 9.5: Torque based Chirp excitation signals for main and idling test

For the pedal based modelling only the torque signal was different, however the resultant variation in loads was similar. For the normal running phase engine speed was varied in the range 1000-2500rpm and engine torque was varied in the range 20-250Nm (pedal equivalent 20-50%). During the idle test engine speed and torque were varied in the ranges 790-1050rpm and 0-25Nm respectively to provide some overlap with the main data set. The three other inputs were controlled via adder functions included in the ECU architecture and varied about their production calibrated values with SOI +/-2°CA, MAF +/-10% and rail pressure +/-100bar. For the idle test these limits were reduced for MAF and rail pressure to +/-5% and +/-25bar respectively.

2.2.2. Capturing thermal dependent behaviour

To establish the thermal behaviour of the engine during warm-up two identical tests were performed, the first from a cold starting and the latter from hot starting engine. The difference in NO_x between the two configurations was then measured to establish the temperature-dependent scaling function. As described earlier in this chapter, this scaling function is expected to be different for torque and pedal based models, and so this experimental procedure was performed twice with different engine load control.

It was important that the only difference between the cold- and hot-start tests was the engine temperature. However, experience from previous work in this thesis has highlighted strong interactions between the thermal state and the control strategy. As a result, engine emissions and fuel consumption during warm-up are affected in two ways: directly by the colder combustion chamber and oil temperatures and indirectly by changes in engine control such as injection timing. From a modelling perspective, the changes in ECU control are accounted for as inputs to the model and should therefore be inhibited in the training data. Because full access to the ECU was available, the major temperature-dependent functions were identified and isolated, meaning any recorded changes in other variables were purely physical.

The training sequences for torque and pedal based models are shown in figures 9.6 and 9.7 respectively. For the torque based model this was simply the NEDC because this test is run in torque set-point mode but for the pedal based model a different signal was prepared based on the chirp inputs. The plots show superimposed data from both the cold starting engine (blue) and the hot starting engine (red).

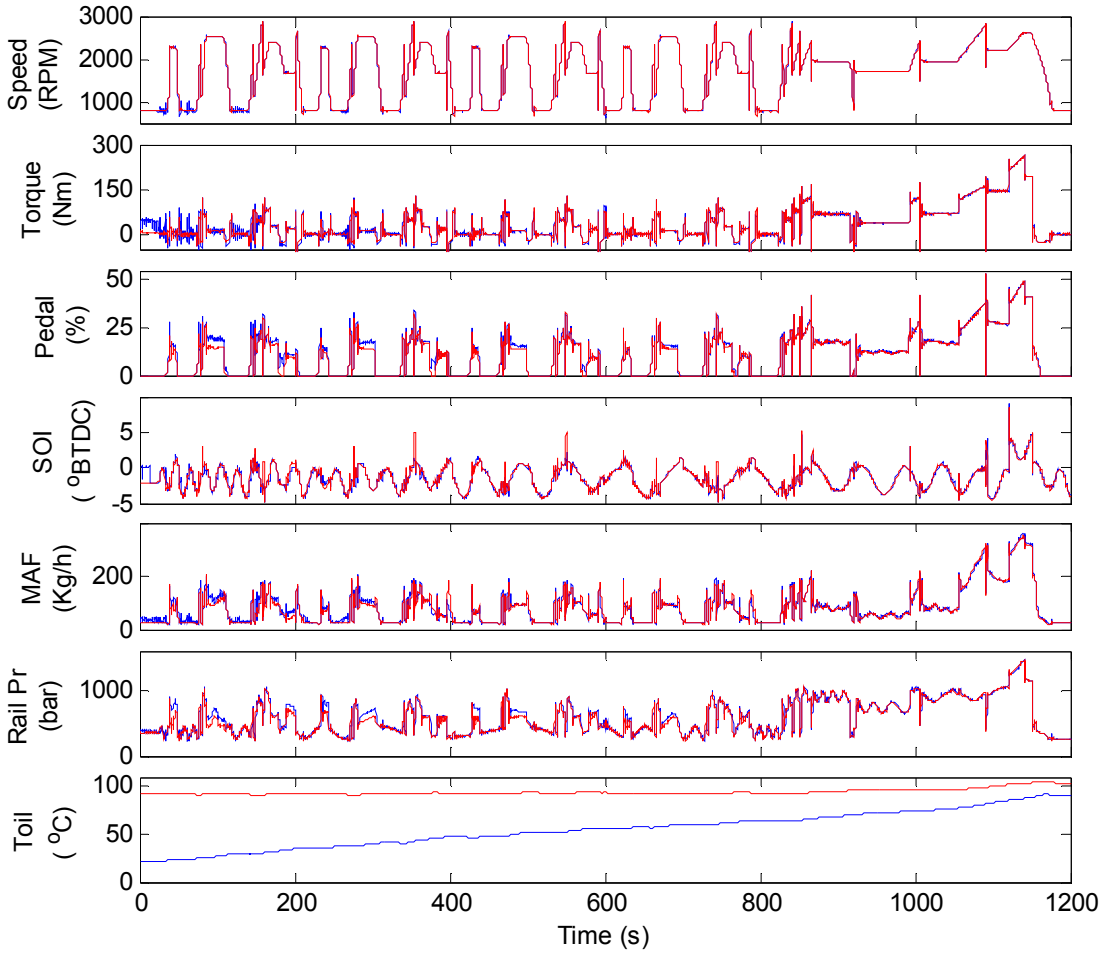


Figure 9.6: Training data for temperature-dependent scaling factor for torque based model

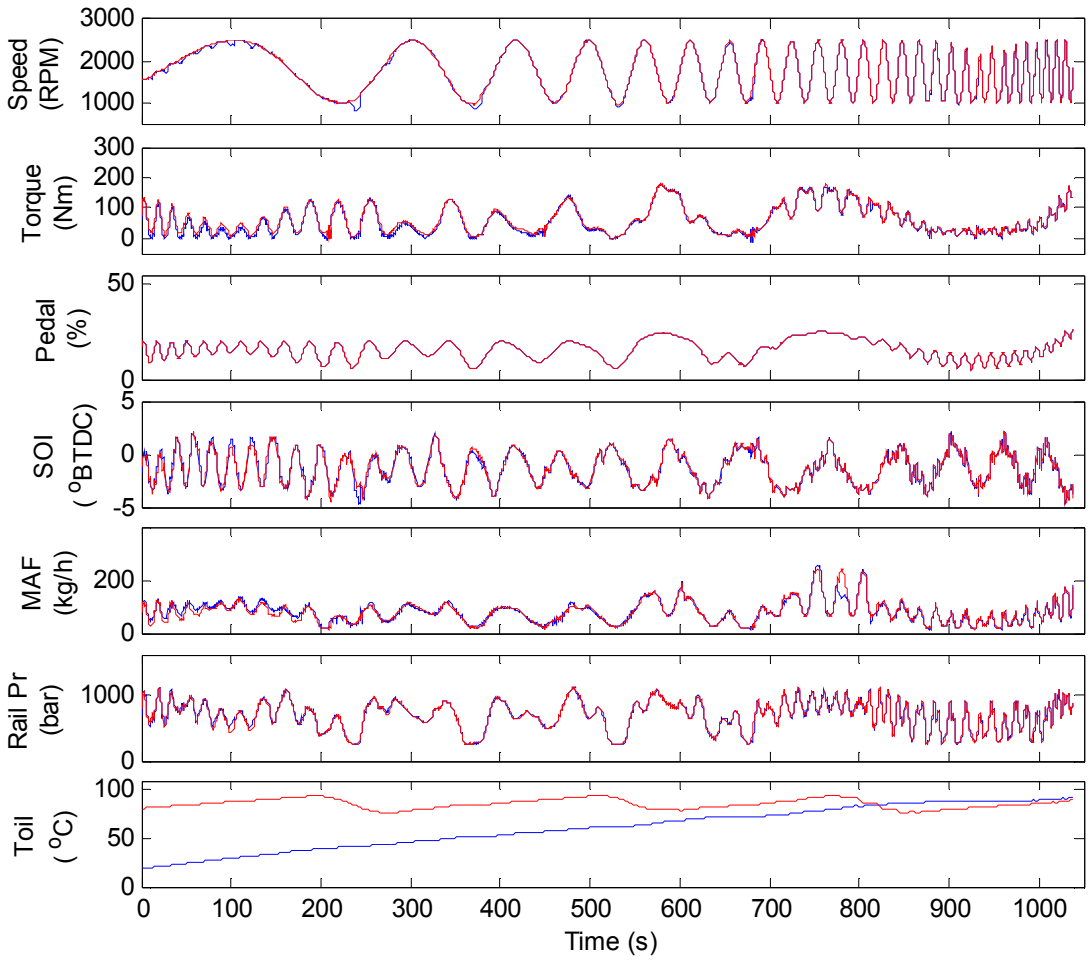


Figure 9.7: Training data for temperature-dependent scaling factor for pedal based model

From these measurements the following observations were made:

- Engine speed is repeatable for both modelling approaches.
- Engine torque is repeatable for the torque based approach as this is the demand signal. The pedal position is slightly higher in the cold test owing to higher fuel required to overcome higher friction (figure 9.6).
- Conversely in the pedal based approach, pedal position is repeatable but measured torque is lower for the cold-start test (figure 9.7).
- The three ECU control variables appear to agree well between the hot and cold-start tests with the largest difference appearing for rail pressure in the torque based test. This is expected to be a result of the changes in pedal position mentioned previously.

3. *Results*

3.1. Training signal and convex hull

Because the excitation of ECU variables was achieved through adder function, the actual design space coverage can be seen after the training sequence had been run. A simple representation of this design space is achieved using a matrix plot: This shows pairwise scatter plots of the inputs and this is shown for the torque based model in figure 9.8. In this figure the normal running phase data is shown in green, the idle phase in blue and the NEDC is superimposed in black. It is immediately apparent that although the majority of NEDC points overlay the training data, there are a large number of points outside the training region, notably apparent on the speed vs. torque plot:

- The points with negative torque represent each of the simulated downshifts; these motored conditions were not replicated in the training data.
- There are some excursions on engine speed above 2500rpm. This was due to unforeseen changes to the experimental facility during the work program and it has been noted that this should be rectified in future work.

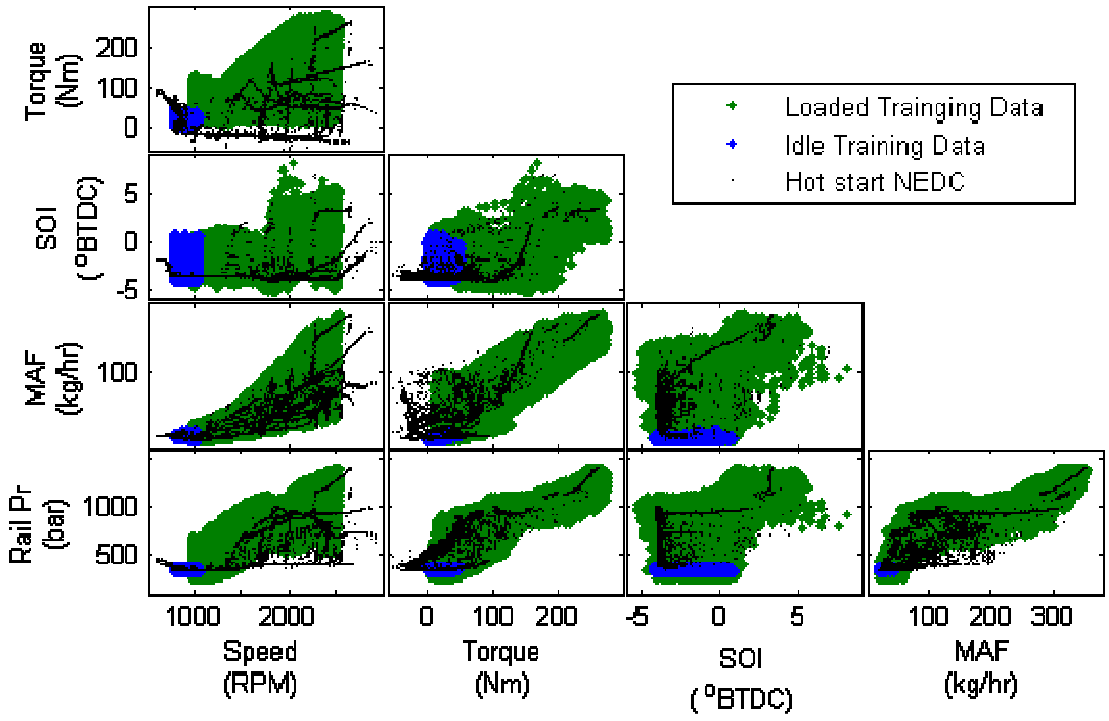


Figure 9.8: 2D projections of the 5D design space showing the main training data, idle training data and NEDC validation data

It is important to note that although the above representation is simple and does not require any further mathematical effort, a point that appears within each the training data in each of the scatter plots may not actually be within the training data region. This is because the matrix plot shows 2 dimensional (2D) projections of the multidimensional input space. To illustrate this point the example in figure 9.9 should be considered: both 2D projections and a 3D plot of an arbitrary training and validation data-set are shown. The training data, plotted in blue, is essentially a 3-level full factorial experimental design, with one of the corner points missing which could arise if this point were not achievable or a measurement error occurred during data acquisition. The validation data, shown in green, is a random uniformly distributed data set within the range the 3 inputs. Considering only the 2D projections, all of the green validation points appear to lie within training data region. However looking at the 3D plot, clearly there are a number of points the lie outside the space created by the training data point. This space is referred to as the hull.

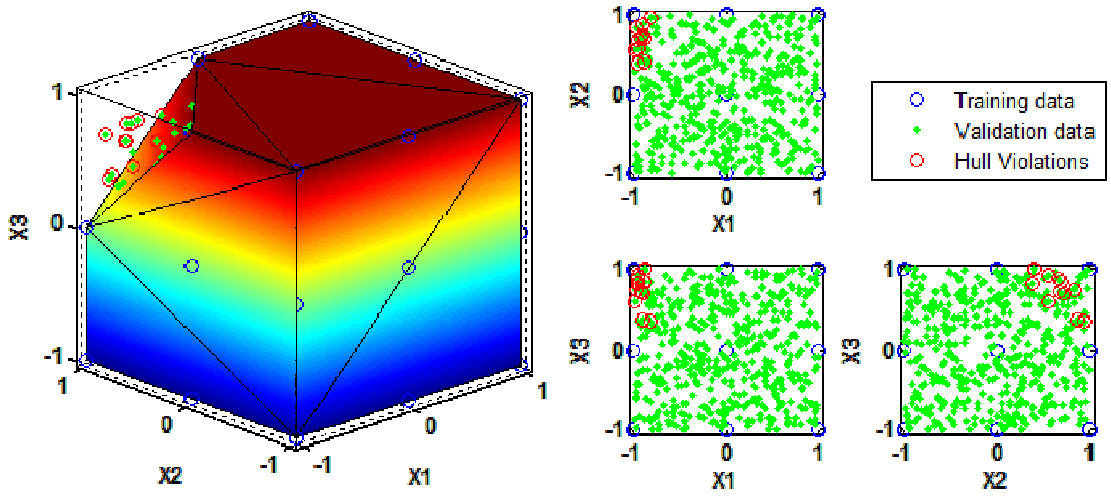


Figure 9.9: Example of 3D hull from arbitrary 3-level full factorial design missing one corner point and random validation data illustrated in 3D and 2D projections. This illustrates that despite all data points appearing to be within the training data region on the 2D projections, there are clearly points outside the training region when viewed in 3D.

The example in figure 9.9 is fairly simple owing to the simple shape and low number of inputs. However the principles applied are the same when applied to the 5D example produced by the chirp signals. The details of the algorithm for identifying hull points is beyond the scope of this thesis but the process is analogous of wrapping an elastic band around a group of nails in a block of wood: the elastic band will define a perimeter based on those nails at the extremity of the group. Some examples of 2D convex hulls are shown in figure 9.10. The algorithm used in this work was the *Qhull* algorithm [157] because this has been used in production software

application for steady state DoE[96] and is available is common Matlab based toolboxes[97]. The convex hull algorithm is not appropriate to all DoE applications, and a typical shortfall is illustrated in the third frame of figure 9.10: the two extreme points cause the hull to cover a large area on the scatter plot but the data is concentrated in the right hand region. This suggests large area of trust region in-which there is no recorded data and no knowledge of system response. Because of the continuous nature of the chirp signals, it was judged this shortfall would not affect this work.

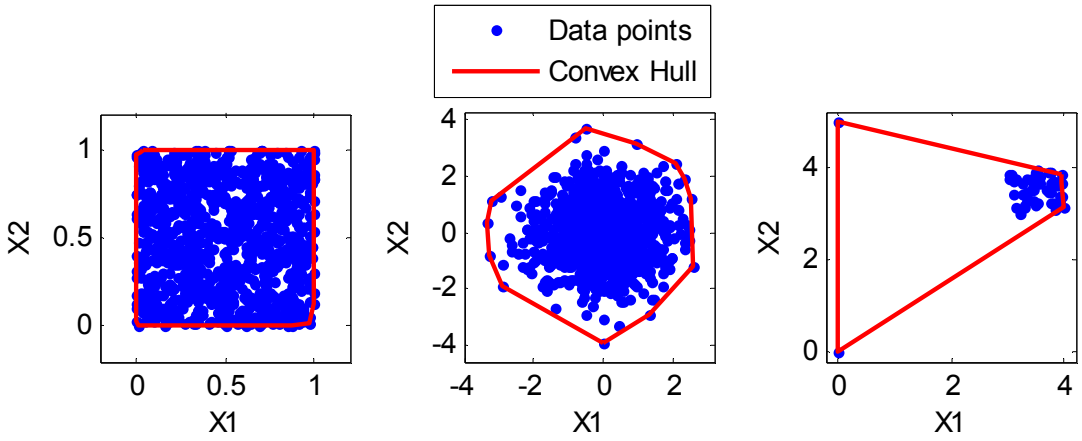


Figure 9.10: Three example convex hulls for different 2D data sets

The hull algorithm returns a set of inequalities relating to the facets of the hull and defining the hull mathematically within the design space. When assessing validation points against the hull, each of these inequalities should be met for a point to be considered within the hull. It is usual to apply a tolerance at the edge of the hull: this accounts for rounding errors it is set to a small value but can also be used to control the degree of extrapolation.

Although this methodology has been applied successfully to static DoE, there are some limitations when applying it to dynamic modelling work. Effectively any point in a validation sequence may statically lie within the training region, but not dynamically. To rigorously account for this the hull dimensions should be augmented into the dynamic region. For example the 5D static hull in this work would become a 10D problem if one delay term per input was used; a 15D with two delays per input and so on. This significantly increases the complexity of the hull and subsequent calculation times. In addition, it was accepted that the hull approach would not be used to precisely identify the exact location of every point with

[96] This is used in IAV EasyDoE commercial software for applying DoE methods in engine calibration.
[97] See Mathworks Matlab software documentation for '*convhulln*' function.

respect to the hull, but rather as a mathematical way of identifying clear outliers that would be affecting the results. It was therefore decided to use a convex hull based solely on the static inputs.

3.2. NO_x model

3.2.1. Global fully-warm dynamic model

The procedure for fitting the Volterra models is similar to the response models calculated in chapter 7, but in addition to the model order and interaction levels, the number and order of delay terms also needs to be chosen. The model parameters are subsequently selected using the OLS approach (see chapter 3). NO_x emissions models performed best when not using any delay terms. This means that despite the dynamic nature of the excitation, NO_x emissions concentration were predictable based on static inputs alone. Details of the model terms are provided in table 9.2.

	Torque based model	Pedal based model
Model order*	4^{th}	4^{th}
Interaction order#	2^{nd}	2^{nd}
Delay terms	None	None
Total Number of terms	17	15
Output transform&	0.25	0.25

* Model order refers to the maximum exponent of main effect terms within the model
# Interaction order refers to the maximum number of main effects considered for any single model parameter
& The output transform is a variance stabilising transform [112] (sometimes referred to as a Box-Cox transform) which aims to achieve a constant variance over the model range and simplify the required model

Table 9.2: Summary of model structures for NO_x response models

An extract showing how the torque based model fitted the training data is shown in figure 9.11 along with the 'fitted against measured' plot for the whole data set. The base polynomial models fitted well with R^2 values for both the pedal and torque based models around 0.92-0.93. When the dynamic filtering is applied to account for the sensor dynamics, the fit improved with R^2 values around 0.94-0.95. The effect of this is clearly shown in the detailed view of figure 9.11 where the data including the sensor dynamics estimates the measured data more closely and acts as a filter on the base polynomials. This resulted with normalised residual errors of about 6% which was higher than some previous studies [119], but still satisfactory.

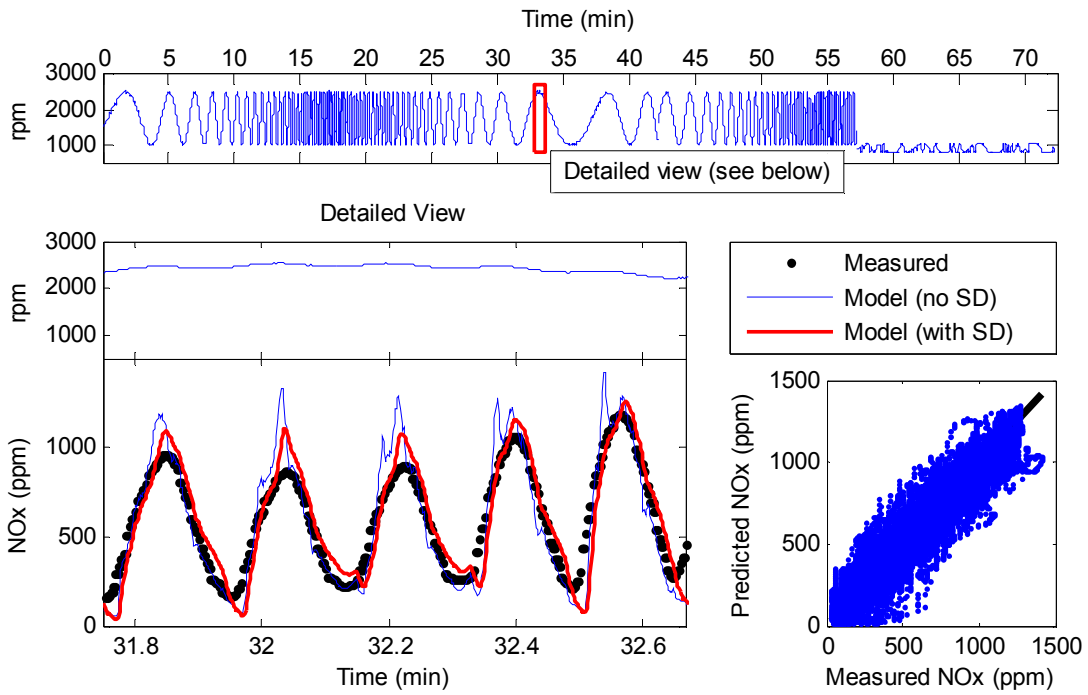


Figure 9.11: Fitted NO_x emissions for training data for torque based model with detailed view for 1min period. The effect of including sensor dynamics (SD) in the model is highlighted. Also shown is the overall fit showing measured vs. fitted plot for complete data set.

Initially, models were fitted without the idle phase training data and unsurprisingly these performed poorly when predicting the idle phases of the NEDC [155]. The inclusion of the idle training data meant the model orders were increased to those detailed in table 9.3. Figures 9.12 and 9.13 show predicted hot-start NO_x emission over the NEDC. In the latter of these figures, points calculated "outside" the convex hull are marked in red on the engine speed trace. The prediction performs well with the exception of four large predicted spikes which exceed 4000ppm. Close inspection shows these occur during the simulated downshift at the end of the 35km/h cruises. These points correspond to high frequency excitation and large negative torques. Figure 9.13 identifies these points as outside the training data hull and will be removed from the data when calculating modelling statistics[98].

[98] Because of the dynamic nature of the modelling, a complete time series is required and these points can't simply be removed from the data. In subsequent model use, some simple correction procedures can be applied to the model after simulation to avoid these issues. For example, NO_x can be set to 0ppm for all conditions where torque is below 0Nm.

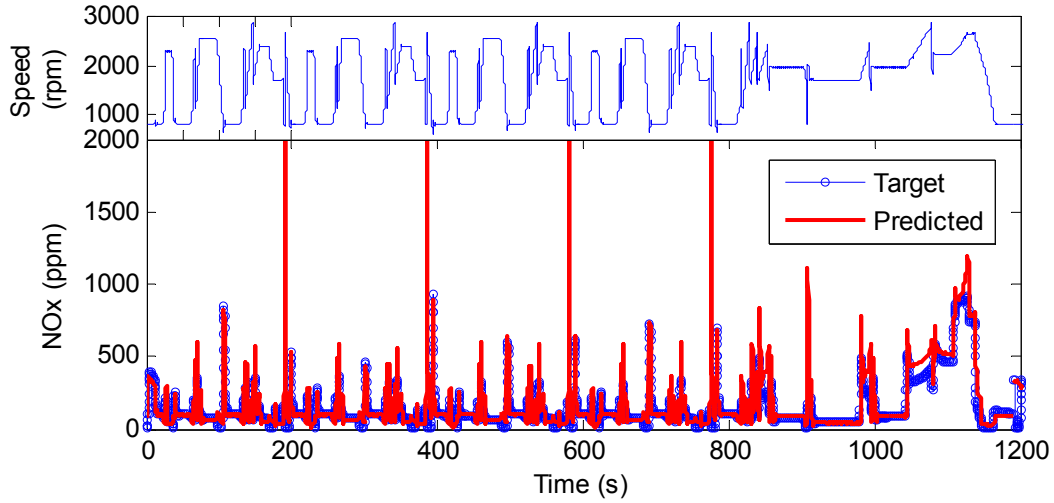


Figure 9.12: Predicted hot-start NEDC NO_x for torque based model

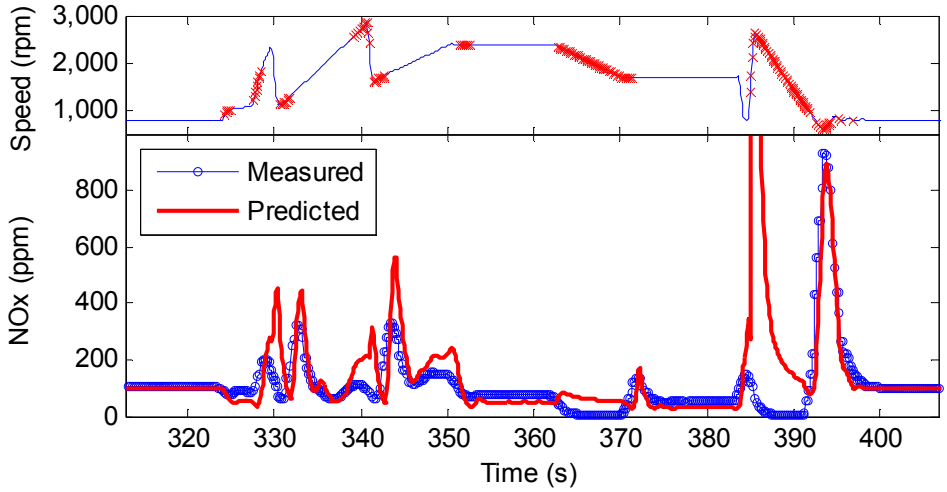


Figure 9.13: Detailed view of predicted hot-start NO_x for torque based model with idle training data highlighting points outside model training hull on engine speed trace.

The hot model was then used to predict NO_x emissions for a cold-start NEDC cycle and the results are shown in figure 9.14 (the four downshifts have been removed for clarity). The model over-predicts at the start of the test but the prediction gradually becomes better as the engine warms up. NO_x emissions would be expected to be lower following a cold start as their formation is strongly linked to combustion temperatures and has been demonstrated throughout this thesis. A similar modelling shortfall was observed in previous work [119] when using fully-warm models to predict behaviour from cold-start. Effectively the model is violating its convex hull in the temperature domain, although this is not explicitly expressed. The model is only valid over the training data region which was achieved using a hot engine.

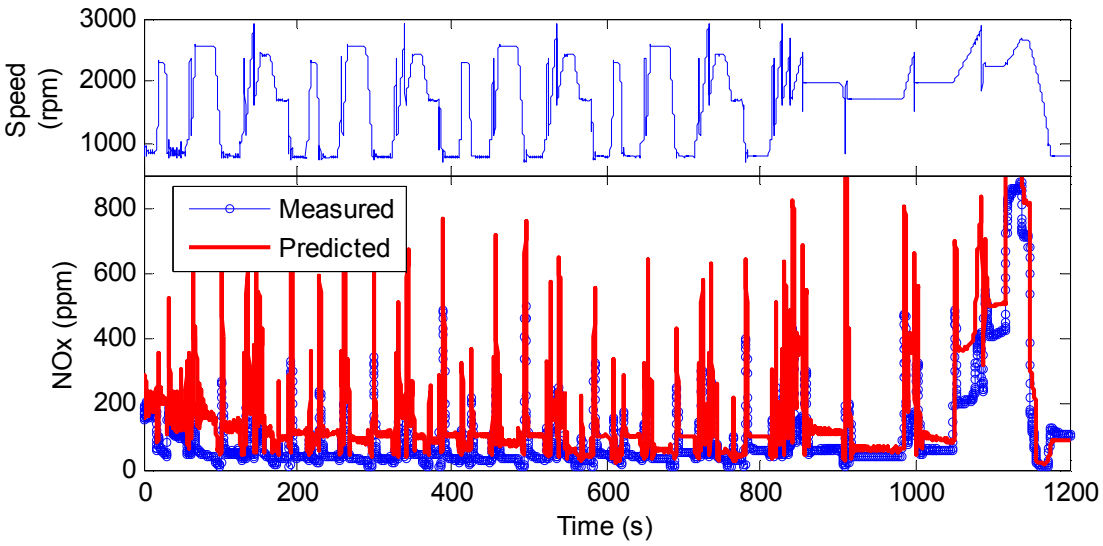


Figure 9.14: Predicted cold-start NEDC NO_x for torque based model (problematic downshifts removed for clarity)

Table 9.3 shows the modelling statistics for the prediction of two independent hot-start NEDC tests. Extrapolation was allowed for some points outside but close to the periphery of the hull by specifying a tolerance limit: this was set to 3% for the torque model and 1% for pedal model.

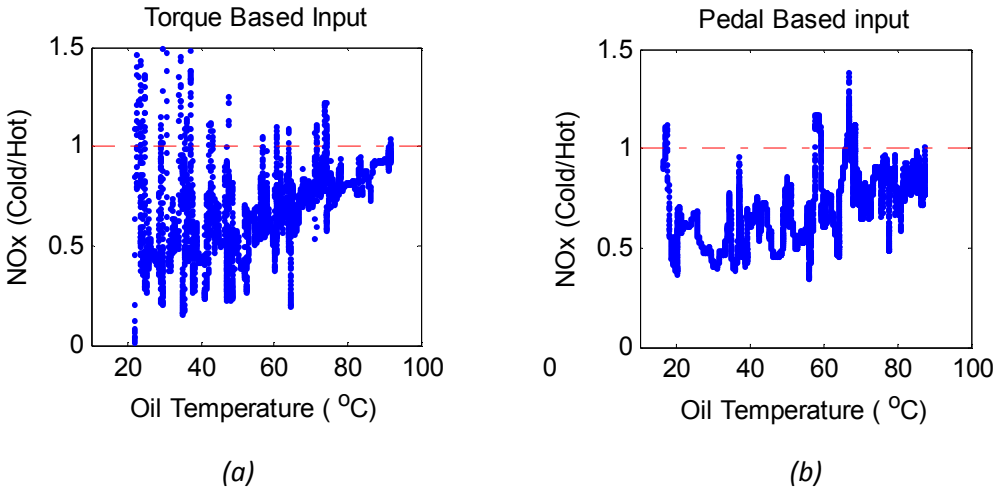
Load control	Torque	Pedal
Hull limit tolerance	3%	1%
% points included using hull	71%	62%
Predicted R^2	0.81	0.78
RMSE (ppm)	65	67
Normalised RMSE (%)	6.2	6.8
Signal to Error ratio (dB)	10.3	9.9

Table 9.3: Modelling statistics for predicted hot-start NEDC

The models performed similarly for hot-start prediction with R^2 typically in the region of 0.8 and RMSE 65ppm (approx. 6.5%). The pedal based model may struggle during idle phases as when the ECU idle controller is working, pedal position will always be 0% whereas torque will vary. This is a shortfall of the pedal based modelling which would probably be solved by using a fuelling based approach.

3.2.2. Thermal model

The measured ratio of cold- to hot-start NO_x emissions is shown in figure 9.15 for both torque and pedal controlled tests. In this case oil temperature has been used because of its effects on piston temperature through piston cooling jets. Because of the low level of instrumentation on the PUMA 2.0L engine, there was little choice for reference temperatures, but in another facility other temperatures could be used. If the model is to subsequently be used in conjunction with a lumped capacity model, most temperatures will be available during simulation. In a calibration task, like this engine very few temperatures are likely to be available.



(a) *(b)*

Figure 9.15: Cold-start to hot-start NO_x emissions ratio for (a) torque and (b) pedal based experiments

The first remark in both cases is that although there is a simple overall trend, the detailed response is complex and appears to be affected by speed and load changes, suggesting significant interactions between these variables which cannot be modelled by the structure proposed in this work. The variations may also be a result of the misalignments between the two cold- and hot-start data sets. Both hypotheses are possible and further investigations would be required to identify the true cause. However, as the overall trend is simple, the modelling approach adopted here should provide some improvement on cold-start prediction. In both cases the overall trend is similar and the cold- to hot-start NO_x ratio is approximately 0.5 at an oil temperature of 20°C and rises to 1 at 90-100°C. The scaling function was obtained by regression to the graphs presented in figure 9.15 and both a linear and quadratic functions were considered.

Figure 9.16 shows the predictive performance over cold-start NEDC. The model performs much better during the cold phase at the start of the drive cycle, but also at the end of the cycle. This is explained by the fact that at the end of the NEDC the engine has still not reached fully warm conditions. The reason warm-up period is extended on the PUMA 2.0L engine is because of the significant additional thermal inertia of the water/water heat exchanger that simulates the radiator.

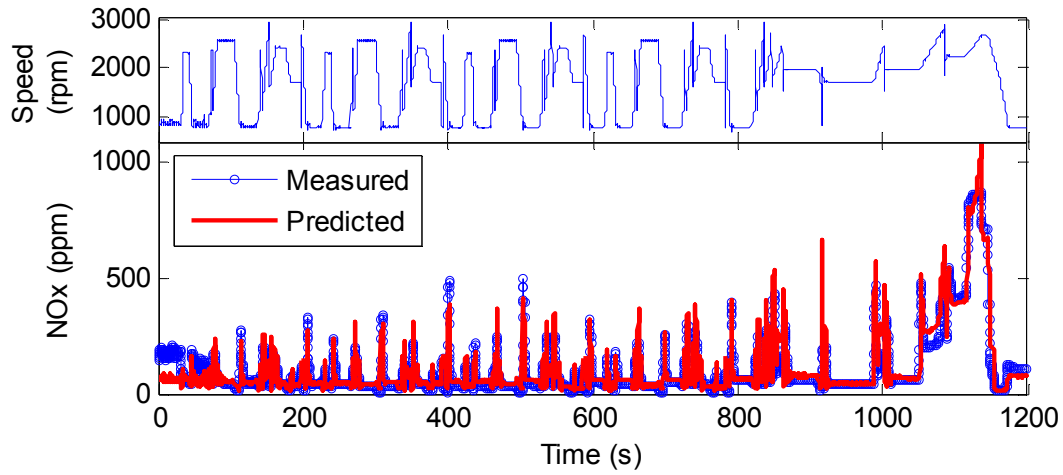
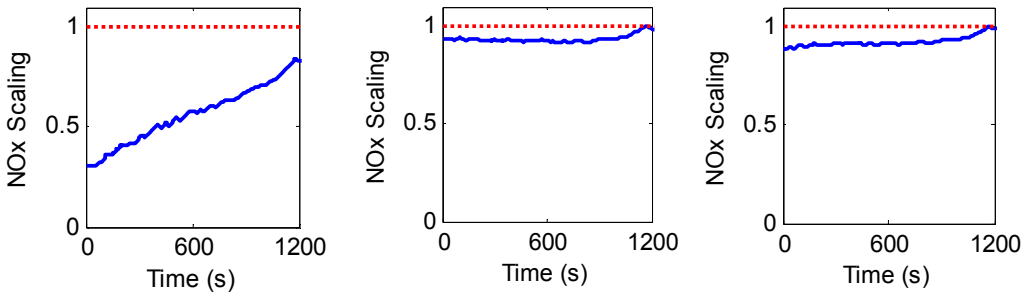


Figure 9.16: Temperature-dependent NO_x model performance for cold-start NEDC for torque based input (problematic gearshifts removed)

Figure 9.17 shows the scaling factors throughout the NEDC tests for both cold-start test and two hot-start tests. This confirms that in the cold-start test the oil temperature did not reach fully warm operating temperature. It is also interesting that the temperature scaling factor is applied mildly to the hot-start NEDC where the low power phase causes a drop in oil temperature.



(a) Cold Start *(b) Hot start 1* *(c) Hot start 2*

Figure 9.17: NO_x scaling function for (a) cold-start and (b), (c) two hot-start validation NEDC tests (red dotted line shows the upper limit of the scaling factor set as 1)

Table 9.4 shows the modelling statistics for predicting NO_x emissions over the cold-start NEDC using different scaling factors. As with the hot-start tests, points calculated "outside" the convex hull were removed before statistical analysis. Both torque and pedal based approaches perform poorly without any scaling factor with R^2 values of 0.63 and 0.68 respectively. In both cases, the inclusion of the scaling factor significantly improves the model performance over the cold-start NEDC, however the torque based model performs better. There is little difference between a linear or quadratic temperature-dependent scaling function. This suggests that although in theory the scaling factor should be more complex for torque based models, in practice this effect is masked by other inaccuracies of the model.

Engine load Mode	**Torque**			**Pedal**		
Scaling	None	*f(T)*	*f(T,T²)*	None	*f(T)*	*f(T,T²)*
Hull limit	3%	3%	3%	1%	1%	1%
% points included	70%	70%	70%	65%	65%	65%
Predicted R^2	0.63	0.9	0.89	0.68	0.75	0.76
RMSE (ppm)	115	42	49	102	66	65
Normalised RMSE (%)	14%	5%	6%	12%	8%	8%
Signal to Error ratio (dB)	4	12	11	4	8	8

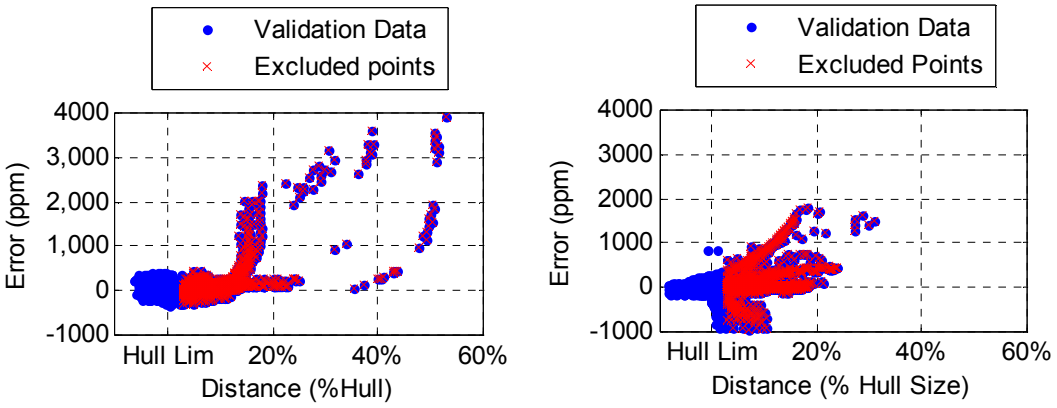
Table 9.4: Predictive NO_x model performance over cold-start NEDC with different scaling factors

3.2.3. Model discussion

For both torque and pedal based modelling approaches, figure 9.18 shows the prediction error against the calculated distance from the hull distance[99]. The errors for points outside the hull will depend on the model structure and in which way the hull limits are violated. If a point is outside the hull for a variable with low order relationship, then there will be a lower chance of high errors, however if the hull limits are violated for a variable with high order response, then there is more chance of extreme predictions outside the design space.

Within the hull (distances lower that 0%) both models perform similarly with fairly small errors, however the behaviour outside the hull is very different for each model. The torque based model is quite stable in extrapolation even up to distance 10% of the hull size. Above this there is an exponential increase in prediction error which manifested itself during the simulated down-shifts. Conversely, the pedal based model exhibits a large increase in error in the near proximity of the hull limit.

[99] The hull distance is expressed in percentage hull size, with 0% representing the hull limit.



(a) Torque Based Model *(b) Pedal Based Model*

Figure 9.18: Assessment of NO_x model performance in extrapolation for (a) torque based and (b) pedal based model. Distance from convex hull against model prediction error for NEDC validation data shows where model error increases rapidly at larger distances from hull

To explain these differences, the example gear shifts will again be used as an example. The lower limits of the 1D hull in engine load dimension will be 0Nm torque and 0% pedal position. In the case of the torque model, the points with negative torque will be progressively further away from the hull as the magnitude of the negative torque increases. For the pedal based data, regardless the negative torque, the pedal position will be 0%, meaning all points in the downshifts will be much closer to the hull and depending on other variables may be on the surface of the hull. Consequently, the error will rise more rapidly in this example and the extrapolation tolerance will need to be tightened for the pedal based approach.

3.3. Fuel consumption and CO_2 model

In the initial stages of the dynamic modelling work it was intended to build response models for engine fuel consumption. However in practice there were significant difficulties in obtaining a reliable dynamic measurement of specific fuel consumption across the engine operating range. Due to the length of the training sequence, it was necessary to use the carbon balance method for fuel consumption estimate. The capacity of the gravimetric fuel beaker was too small and refilling would be necessary during the experiment causing a discontinuity in gravimetric measurement[100]. Even if the gravimetric result were available, both measurement principles provide an estimate of fuel flow; to get a measure of specific fuel

[100] The fuel balance had a fuel capacity of 1L which was sufficient for NEDC experiments but not for the chirp excitations despite the hour schedule being split into 20min periods because of more prolonged operation at higher engine power. More details on the gravimetric fuel consumption measurement are discussed in chapter 3.

consumption this needs to be divided by brake engine power. At low load such as large portions of phase 1 of NEDC, both the power and fuel flow are small, meaning specific fuel consumption becomes ill-conditioned. The non-perfect time alignment of various dynamic measurements results in the data-set being unusable. As a result it was decided to model CO_2 emissions using the same approach as for NO_x detailed above.

Before building models for CO_2, it is also worth discussing the differences between the torque and pedal based model approaches, notably referring to the thermal based modelling. As highlighted previously, for a torque based model simulating an NEDC, the torque input will be the same regardless of engine temperature because the input relates to brake torque; hence all other inputs being equal, a temperature scaling function would simply increase CO_2 to account for increased friction. Conversely, for a pedal based approach under colder conditions a higher pedal position would be required to produce the same brake torque as under hot-start conditions. This means that the changes in CO_2 under cold conditions will be accounted for in part by a change of pedal position in the fully warm model in addition to the temperature scaling function.

From a training perspective for the pedal based model, both the hot and cold-start tests aimed at deriving the scaling factor will have similar CO_2 emissions as the difference between the two tests will primarily be a drop in brake torque. As a result, the approach of repeated hot and cold-start tests to capture differences in CO_2 in theory does not work when using a direct pedal based strategy. There is still a practical interest in generating a fully warm model for CO_2, and in the following paragraphs the data will be presented to see how this effect appears in practice.

The models were built using the same process described for NO_x. In this case there was a clear benefit from the inclusion of delay terms: the fit R^2 was increased from 0.77 to 0.9 for both models. The delay terms used for the torque and pedal models were 0.9s and 0.7s respectively. The model structures for both the torque and pedal based approaches are shown in table 9.5.

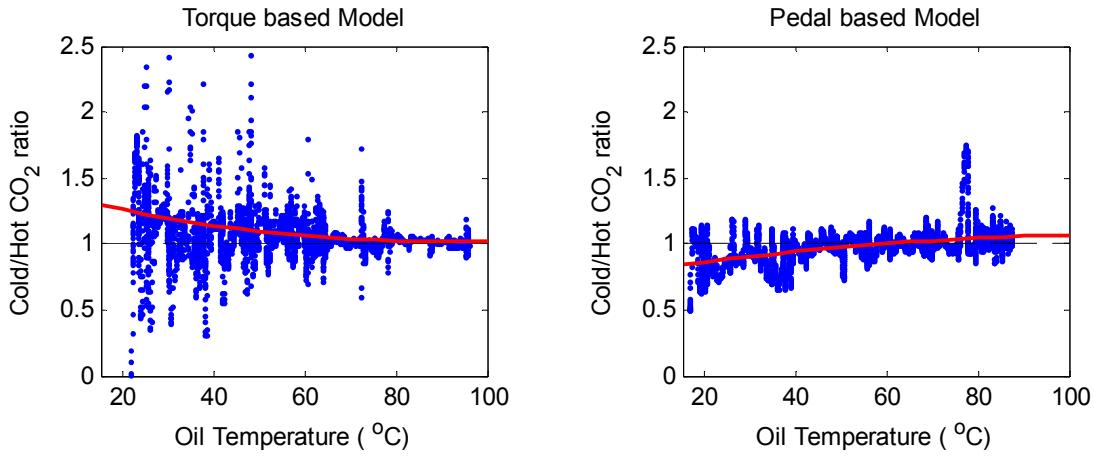
Figure 9.19 shows the ratio of cold-start to hot-start CO_2 for both the torque controlled and direct pedal set-point experiments. The first point to note is that the effect of temperature on

this ratio for CO_2 is smaller than for the same ratio of NO_x. This is because the scope to reduce CO_2 emissions by faster warm-up is small relative to overall CO_2 emissions.

	Torque based model	Pedal based model
Model order*	4th	3rd
Interaction order#	2nd	2nd
Delay terms$	1 (0.9s)	1 (0.7s)
Total Number of terms	27	15
Output transform&	0.75	0.75

* Model order refers to the maximum exponent of main effect terms within the model.
\# Interaction order refers to the maximum number of main effects considered for any single model parameter.
$ Both the number of delay terms and the time delays are specified.
& The output transform is a variance stabilising transform [112] (sometimes referred to as a Box-Cox transform) which aims to achieve a constant variance over the model range and simplify the required model.

Table 9.5: Summary of model structures for NO_x response models



(a) *(b)*

Figure 9.19: Cold/hot CO_2 ratio for (a) torque and (b) pedal controlled experiments used for deriving temperature scaling function

The shape of the data fit is as expected for the torque based experiment with colder engine temperature causing higher CO_2. As the temperature rises, the ratio reduces from 1.25 at 20°C to 1 above 90°C. Strangely, the opposite effect is seen with the pedal controlled experiment: there are in fact slightly reduced CO_2 emissions at colder temperatures. On close inspection of the training data, this appears to be a result of small differences in mass air flow under the colder conditions. Although care was taken to ensure minimal variations in control parameters between the cold and hot experiments, clearly this has not been achieved in this case.

Model fit statistics for validation NEDC cycles are shown in table 9.6: these results are shown for both the torque and pedal based approaches, for a hot NEDC and cold-start NEDC and with and without temperature scaling factor.

Engine load Mode	**Torque**			**Pedal**		
Scaling	Hot	*Cold No scaling*	*$f(T,T^2)$*	Hot	*Cold No scaling*	*$f(T)$*
Hull limit	1%	1%		1%	1%	
% points included	65%	65%		65%	65%	
Predicted R^2	0.82	0.92	0.92	0.81	0.84	0.86
RMSE (%CO_2)	1.22	0.7	0.86	1.33	1.2	1.0
Normalised RMSE (%)	11.3	5.9	7.3	12%	10	8.8
Signal to Error ratio (dB)	14	19	18	13	14	15

Table 9.6: Predictive CO_2 model performance over cold-start NEDC with different scaling factors

In both torque and pedal base approaches, the model fits the cold-start drive cycle better than the hot-start which remains unexplained. There is also very little improvement to model fit when using the temperature-dependent scaling function:

- For the torque based NO_x model, the inclusion of the temperature scaling factor improved the cold-start validation data fit from an R^2 value of 0.63 to 0.9.
- For the torque based CO_2 model, there is little difference from using the temperature based scaling function, with no change in R^2 value.

This behaviour is best explained by the relative impact of engine temperature on CO_2 and NO_x relative to overall emissions levels over the drive cycle. From the work conducted in the previous chapters, the difference in CO_2 over a hot and cold-start NEDC is approximately 100g or 3%; the same number for NO_x emission is 1g or 18%. This difference means the modelling requirement for temperature scaling function for CO_2 is more demanding and trying to identify significantly smaller differences. Unfortunately these differences will become closer to the accuracy of the hot model.

4. *Discussion*

The models produced by the analysis in this chapter allow the prediction of feed-gas emissions concentration over a given duty cycle. For engine manufacturers the accuracy of cumulative emissions over the duty cycle is of key importance. These numbers are easy to calculate using the air mass flow rate and the procedure prescribed by British Standards (as discussed in chapter 3) [127].

In this work, MAF is an input to the model. However, in most practical application, the MAF would not be known and is effectively linked to engine operating point. If the excitation signals used for modelling had been applied to the actuators directly (speed, fuelling quantity, EGR valve position and VGT position), then it would be possible to model air flow rate in the same way as emissions modelling. This estimate of MAF could then be used both in the models calculated here and to calculate overall emissions by mass. This is the method applied by other authors when more development time is available on the experimental facility with possible implementation of prototype control methods [119].

Figure 9.20 shows the measured and predicted cumulative CO_2 and NO_x emissions over a cold-start NEDC. There is a clear offset between measured and predicted values which is a result of the compounded errors throughout the cycle. To minimise the impact of the prediction spikes during simulated down-shifts (see figure 9.12), a simple post-simulation adaption of the model was implemented. This effectively set emissions to 0% or 0ppm in the case of negative torque. Final emissions predictions agree with the validation data to within 6% and 12% for CO_2 and NO_x respectively.

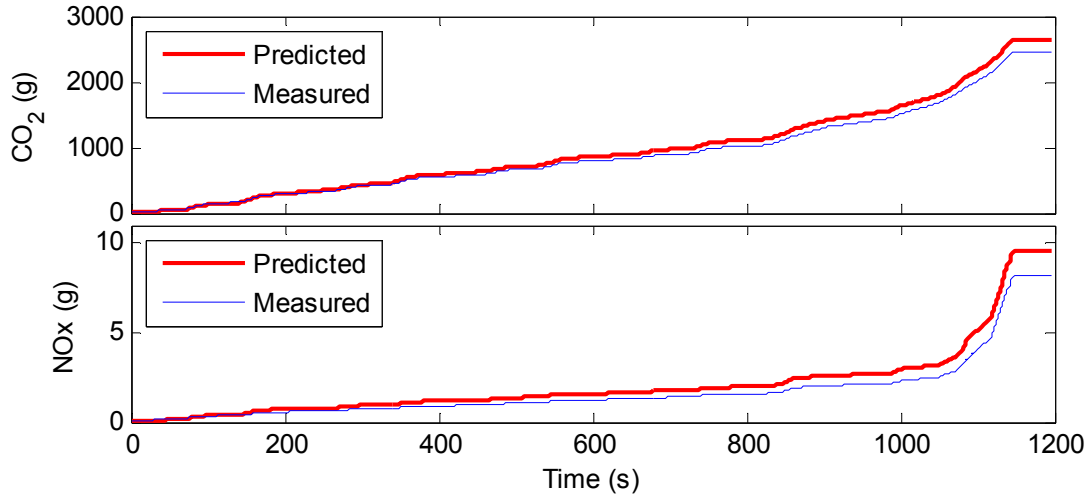


Figure 9.20: Performance of CO_2 and NO_x model for predicting cumulative NEDC emissions

With a validated model for engine emissions, it is possible to integrate their capability into a complete tool for engine behaviour prediction. Changes to engine warm-up rate can be simulated using lumped capacity models that were discussed in the literature review in chapter 2. When validated correctly, these can provide appropriate engine temperatures that can be used in this model. Engine speed and load can be given by a prescribed duty cycle while SOI, MAF and rail pressure can be adjusted by simulating the engine calibration software. The

inclusion of each of these aspects is beyond the scope of this thesis. However to illustrate the potential use of this tool, some modified warm-up rates have been calculated[101] to simulate possible outputs from a lumped capacity model. These modified warm-up rates are shown in figure 9.21. These represent somewhat ambitious warm-up improvements in light of results from previous chapters, however they could be considered as an aggressive reduction in thermal inertia.

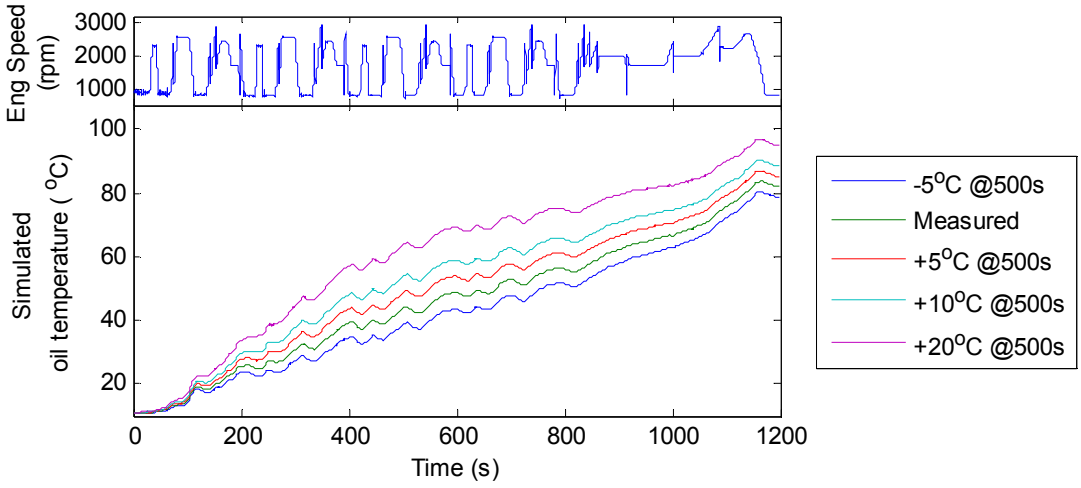


Figure 9.21: Simulated oil temperatures based on an offset from measured engine warm-up rate

A number of simulations have been performed using the different warm-up rates and changes in injection timing. The final cumulative values for NO_x and CO_2 form a trade-off that is illustrated in figure 9.22. The changes in injection timing are simple and applied grossly over the whole drive cycle, however more detailed refinement could be applied to an optimisation problem.

The trade-off is of similar shape to that presented by Brace et al. [23] for varying operating temperatures under steady state conditions. Over the simulated NEDC, changes in warm-up rate appear to have a more favourable impact on the trade-off than changes in injection timing. This was the case a lower load in the results published by Brace *et al.* As the majority of the NEDC cycle is low load, these results would be expected to agree with lower load steady state results.

[101] The modified temperature profiles were obtained through addition or subtraction of a skewed bell-shaped curve to simulate faster warm-up followed by convergence as the engine reached fully-warm conditions.

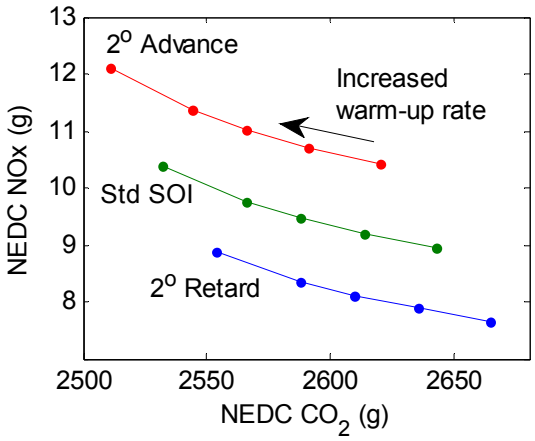


Figure 9.22: NO_x/CO_2 trade-off resulting from simulated NEDC at varying injection timing and warm-up rate

5. *Chapter summary and conclusions*

In this chapter global dynamic models have been produced for NO_x and CO_2 emissions and their performance has been proven over the NEDC. Polynomial based modelling approaches from the literature were reproduced at the University of Bath with acceptable results. These models were subsequently enhanced to include warm-up effects and the improvement was measured based on prediction performance of a cold-start NEDC. Possible uses of these models in conjunction with other simulation tools was then demonstrated by producing NO_x/CO_2 trades-offs for thermal and injection timing changes. The conclusions from this chapter are listed below:

1) Polynomial based modelling approaches are useful tools for the dynamic modelling of NO_x and CO_2 emissions. The modelling of specific fuel consumption was not possible because of difficulties in obtaining an accurate dynamic measurement.

2) The use of a temperature dependent scaling function improved the modelling of cold-start behaviour. Over the NEDC, prediction RMSE was reduced from 115ppm to 42ppm for NO_x and from 1.22% to 0.7% for CO_2.

3) In theory pedal based modelling should provide simpler thermal response but in practice these differences are masked within global model inaccuracies. Torque based modelling provided better quality models then their pedal based counterparts.

4) The modelling approach also performed well for cumulative predictions however model control using multi-dimensional convex hull is required. If this is not used then large compounded errors can result over the 20min simulation.

5) The dynamic models can be used in conjunction with lumped capacity models to provide information relating to design changes in the simulation environment.

Chapter 10 - *Conclusions and outlook*

In this final chapter an overview of the work and achievements is presented. The chapter is organised into three sections:

- *A brief summary of the project work.*
- *The main conclusions alongside the project objectives that were listed in chapter 1, and ultimately against the project aim.*
- *A brief discussion on the wider impact of the project and its outcomes. This will include scope for further work, some of which is on-going at the time of publishing this thesis.*

1. *Summary*

The work in this thesis has focused on understanding and optimising diesel engine warm-up. In chapter 2, a review of the literature was presented focusing on interactions between the thermal, lubricant and combustion systems. In chapter 3, experimental methods were presented to demonstrate small differences for the optimisation of mutli-dimensional problems. Together chapters 4 to 7 present a methodology for engine thermal management optimisation. Initially candidate hardware was selected based on results from previous studies. Chapters 5 and 6 presented scoping experiments to understand the fundamental operation of the prototype auxiliary systems. This knowledge was subsequently used for a rigorous test design to optimise the system during warm-up. This procedure resulted in a 3% fuel consumption reduction at iso-NO_x conditions. In chapter 8 an exhaust gas heat exchanger was introduced to the system in an attempt to further improve warm-up performance, but ultimately the impact was too small for any significant benefits to appear.

Finally in chapter 9, a global dynamic modelling approach was successfully implemented and enhanced to capture engine warm-up behaviour. It was subsequently shown how this approach could be used in conjunction with commercially available simulation tools for engine optimisation.

2. *Conclusions*

The conclusions from this work will be presented against the objectives laid out in chapter 1. Each of these objectives is reminded below with the respective concluding remarks. Finally the main outcomes of this work will be summarised in line with the project aim.

1. *"Review the literature relating to engine cooling, lubrication and combustion focusing on interactions between the systems."*

This literature review was presented in chapter 2. Both the cooling and lubricant systems are vital to engine operation but there is a clear scope to reduce their energy consumption through the use of active control. Reductions in fuel consumption of the order of 3.5% have been demonstrated. During warm-up, significant reductions in fuel consumption can be achieved by reduced engine friction through faster warm-up. Engine oil viscosity reduces with

temperature which in turn reduces the dominating hydrodynamic friction. Reduced warm-up times are achieved by increasing the available heat energy per unit thermal mass either through reduced thermal inertia or additional heat input. The maximum potential fuel consumption benefit is about 4%. Changes to the thermal management system also impact the combustion event and notably emissions formation. NO_x emissions are strongly linked to engine temperature, but this can be exploited for further benefits if an active thermal management system is included in the calibration procedure.

2. *"Establish experimental methods (hardware and procedures) to confidently demonstrate small changes in fuel consumption to prove cost effectiveness of candidate prototype hardware."*

Both experimental methods and facilities were enhanced over the course of this work. The Design of experiments approach was adapted to capture dynamic behaviour during warm-up. Firstly the conventional DoE approach was used but steady state points were replaced by cold-start drive cycles. Secondly, the dynamic DoE approach was enhanced to capture engine warm-up behaviour. The measurement of fuel consumption can be improved by the inclusion of simple correction factors to account for inaccuracies arising from integration of measurement devices into the experimental system. NO_x emissions were seen to be strongly affected by ambient humidity, more than is suggested by a well-known correction factor and a review of these is recommended. For the purposes of this study, an empirically derived correction factor was proposed to allow direct comparisons of results from experiments with significantly different ambient conditions. The resulting system was capable of confidently demonstrating the expected differences from the literature review within reasonable time frames. In addition, extensive instrumentation was installed on the engine to measure small changes in the thermal state that may not be obvious from macroscopic measurements.

3. *"Oversee the integration of prototype hardware into a production engine."*

Prototype hardware was chosen and integrated based on the findings from the literature review. The devices transformed a passive production system into an active system with the aim of modulating heat flows during warm-up. Although the inclusion of these devices caused

an increase in thermal inertia for the prototype system, a production system could avoid this increase through changes of engine design.

4. *"Develop an understanding of the fundamental engine behaviour with changes to thermal system control during engine warm-up."*

The active thermal management system allowed the engine to warm-up faster because of reduced thermal inertia from isolating the front end of the cooling circuit. With these changes, the coolant could be 10^{o}C hotter at the end of NEDC phase 1. This faster warm-up interacted with the engine control strategy for injection timing which is strongly linked to engine temperature. The engine retarded sooner with faster warm-up which compromised the overall benefit in fuel consumption from reduced friction. This highlighted that in optimising the active thermal management system, the engine combustion control should be included.

5. *"Develop an understanding of the fundamental engine behaviour with changes to lubricant system control during engine warm-up."*

In the production engine the oil filter and cooler present a significant thermal inertia. Heat addition in the external circuit either through oil-cooled EGR or oil/coolant heat exchanger can overcome this inertia. Over the first 200 seconds, although the additional heat is seen in the main gallery, this is not seen in the main bearings. However later in the cycle these are observed. Variations in oil pump displacement also have an impact on oil temperature, which is reflected in the combustion chamber temperature and a change of up to 5% in emissions. However the reduced pumping work can reduce fuel consumption by up to 4% which, relative to the scope of fuel consumption gains, is a considerably larger impact.

6. *"Using the knowledge acquired in objectives (4) and (5), conduct an experiment to provide a detailed understanding of the effects and interactions of the prototype hardware. Use this to subsequently optimise the system calibration during warm up."*

Based on the results from the previous objectives, the optimisation of the active thermal management system would include coolant flow control, dual EGR system and injection timing. Oil flow would be set to minimum flow for maximum fuel consumption benefits. The modelling approach provided suitable response models for most measures of engine behaviour. The isolation of front-end coolant reducing thermal inertia increased warm-up throughout the entire engine by about 6^{o}C at the end of phase 1. This change reduced fuel consumption and increased NO_x by 2% and 5% respectively. Variations in the external oil circuit resulted in a trade-off between upper and lower engine of 6-8^{o}C. A significant secondary effect of oil-cooled EGR was more effective EGR gas cooling which resulted in 5-6% NO_x reduction. Optimised calibrations made use of these benefits to offer fuel consumption benefits at iso-NOx conditions of 3.2% and 2% over phase 1 and 2 respectively.

An exhaust gas heat exchanger was added to the system in an attempt to improve warm-up further through heat addition. Despite high heat exchanger effectiveness, the constraints inflicted by the after-treatment devices caused the additional inertia from device integration to nullify most of the system benefit. Only a small improvement in upper/lower engine warm-up trade-off was measured with no observable impact on fuel consumption and emissions.

7. *"Demonstrate suitable system behaviour under fully warm conditions."*

The optimised thermal management system demonstrated adequate performance under fully warm conditions. Engine-out coolant flow control allowed the engine to operate 8-10^{o}C hotter under part load conditions giving 1% FC benefit but 5.5% NO_x penalty. Oil-cooled EGR did not have adverse effects on EGR gas cooling, provided it is implemented in conjunction with a variable flow oil pump.

8. *"Improve state of the art engine dynamic modelling methods to capture engine behaviour during warm-up."*

The dynamic modelling approach could be used to predict NEDC NO_x and CO_2 emissions, but outputs such as specific fuel consumption could not be predicted because of issues with accurate dynamic measurement. This modelling approach, which had previously only been used for fully-warm conditions, can be enhanced through a simple temperature-dependent

scaling function. The models also performed well in predicting total drive cycle emissions despite effects of compounded errors over the 20min simulation. These validated dynamic models can subsequently be used as part of a warm-up simulation environment in conjunction with lumped capacity thermal models.

The aim of this project was to demonstrate improvements to Diesel engine efficiency through reductions in parasitic losses, notably during warm-up. A number of devices were integrated into a production engine cooling circuit. With the exception of oil-cooled EGR, each of these systems had previously been used individually and published results have highlighted the benefits of these devices. In this work a more global optimisation approach has been taken giving further benefits through the interactions between systems. Through this work, two design of experiments methodologies have been presented for thermal management optimisation: the first through drive cycle experiments with available candidate hardware, and the latter by capturing dynamic engine behaviour during warm-up for use in the simulation environment.

In addition to demonstrating these methodologies that can be applied to future calibration problems, the fuel consumption and NO_x emissions correction factors should be highlighted as they will be applicable and beneficial to future projects both at the University of Bath and other experimental facilities. The fuel consumption correction factor improves gravimetric measurement accuracy during thermal transients where errors of up to 1% were identified. The NO_x correction factor accounted for changes in ambient humidity and was required because a commonly used approach was not aggressive enough for the engine in this study. The empirically derived factor is easily applicable and provides a correction of approximately 2% per g/kg dry air change in humidity.

3. *Outlook and further work*

The work conducted in the context of this thesis has had a number of impacts on other projects. These aspects are listed below:

1) The work related to accuracy and repeatability has been rolled out across the Powertrain and Vehicle Research centre (PVRC) at the University of Bath. A number of software tools have also been developed to introduce these measures without significant impacts on day-to-day work load. In some case, workload has even been reduced. The transferral of these tools to commercial organisations is on-going in the context of a *Knowledge transfer partnership.*

2) It is suggested that the dynamic models established in chapter 9 could be used in conjunction with thermal lumped capacity models to create a global simulation environment for use during engine development. The difficulty in creating these lumped capacity models lies in calibrating the individual masses and heat transfer to give realistic temperature predictions. Often these are validated on a handful of temperature measurements usually relating to the oil end coolant. The experimental work performed in this thesis has yielded an abundance of engine temperatures from many locations in the engine and under many different operating conditions. These measurements could easily be mapped onto individual nodes of the lumped capacity model and provide an in depth validation. At the time of publishing this thesis, such an exercise is on-going at the University of Nottingham.

Other suggested future projects issued from this work that are not currently under investigation within the PVRC but that would merit further contributions are listed below:

1) Further analysis of effect of humidity on combustion is warranted. In a first instance this would revise the correction factor for NO_x emissions that was not adequate following the variations in humidity seen over this project. This could also establish correction factors for fuel consumption and other emissions. A combustion air conditioning unit could be used for accurate control of humidity.

2) The candidate hardware was installed as-best-as-possible in the context of the prototype nature of the project. There may be discrepancies between the performance measured here and the performance in a production calibration. The most important aspect in the author's view is the configuration of the oil-cooled EGR gas loop which may be in part responsible for the more effective EGR gas cooling. It would be useful to produce a single oil-cooled EGR setup to validate this behaviour.

3) Although disappointing in terms of measured benefits, the results from chapter 8 relating to the exhaust gas heat exchanger do form the requirements for a successful system. If resources are applied to the design of a low inertia circuit then significant benefits from this system should be achieved.

4) Integration of the dynamic models established in chapter 9 into a full simulation package is a clear succession to this work. A validated lumped capacity thermal model is required and ideally some simulation of the engine strategy.

References

[1] **N. L. J. Gense, N. Jackson and Z. Samaras**, *Euro 5 technologies and costs for light-duty vehicles. The expert panels summary of stake holders responses*, TNO report 05.OR.VM.032.1/NG, Delft (Netherlands), 2005.

[2] **European commission**, 2008. *COMMISION REGULATION (EC) No 692/2008 implementing and amending Regulation (EC) No 715/2007 of the European Parliament and of the Council on type-approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information.*

[3] **Institute of Mechanical Engineers** (2009), *Low Carbon Vehicles, Driving The UK's Transport Revolution*, London, UK, available from: http://www.imeche.org/knowledge/themes/transport/low-carbon-transport/low-carbon-vehicles (Accessed May 2009)

[4] **European Commission,** 2009, *Reducing CO2 emissions from light-duty vehicles,* Available from: http://ec.europa.eu/environment/air/transport/co2/co2_home.htm (Accessed 5.10.2009)

[5] **Delphi**, 2010. *Worldwide Emissions Standards: Passenger Cars and Light Duty Vehicles*, Available from: http://www.delphi.com/pdf/emissions/Delphi_PC.pdf (Accessed 15.07.2010)

[6] **N. Owen and N. Jackson**, *A new look at the low carbon Roadmap*, presented at the Low-Carbon Vehicles 2009, London (UK), 2009.

[7] **P. Werner, J. Schommers, U. Engel, C. Spengel, C. Reckzugel, M. Paule, T. Maderstein and W. Eisler**, *The New V6-Diesel Engine from Mercedes-Benz*, presented at the 19. Aachener Kolloquium Fahrzeug- und Motortechnik 2010 (19th Aachen colloquium), Aachen, Germany, 2010.

[8] **F.-T. Metzner, J. Hadler, B. Stiebels, T. Jaeckel and L. Caesar**, *Innovative Temperature Management as used in the new volkswagen Touareg*, presented at the 19. Aachener Kolloquium Fahrzeug- und Motortechnik 2010 (19th Aachen colloquium), Aachen, Germany, 2010.

[9] **R. Bauder, J. Kahrstedt, S. WZulch, A. Frohlich, C. Streng, C. Eiglmeier and R. Riegger**, *The second generation 3.0l V6TDI from Audi - the systematic evolution of an efficient engine*, presented at the 19. Aachener Kolloquium Fahrzeug- und Motortechnik 2010 (19th Aachen colloquium), Aachen, Germany, 2010.

[10] **F. Steinparzer, H. Unger, T. Bruner and D. Kannenberg**, *The new BMW 2.0 litre 4-cylinder S.I. engine with twin Power Turbo Technology*, presented at the 32. Internationales Wiener Motorsymposium (32nd Vienna Symposium), Vienna, 2011.

[11] **T. Korfer, T. Schnorbus, M. Klalenborn, A. Kolbeck, G. Bourgoin, M. Cecur and E. Raimondi**, *Integrated Diesel engine concept for lowest CO2-emission requirements*, presented at the 19. Aachener Kolloquium Fahrzeug- und Motortechnik 2010 (19th Aachen colloquium), Aachen, Germany, 2010.

[12] **E. Blanchard, J. Visconti, P. Coblence, F. Legrand, F. Gautier, M. Chevrot, M. Clautet and F. Trochu**, *The New Renault dCi 130 1.6l Diesel engine*, presented at the 19. Aachener Kolloquium Fahrzeug- und Motortechnik 2010 (19th Aachen colloquium), Aachen, Germany, 2010.

[13] **T. Heiduk, R. Dornhofer, A. Eiser, M. Grigo, A. Pelzer, and R. Wurms**, *The new generation of the R4 TFSI engine from Audi*, presented at the 32. Internationales Wiener Motorsymposium (32nd Vienna Symposium), Vienna, 2011.

[14] **R. D. Burke, C. J. Brace, J. G. Hawley and I. Pegg**, *Review of the systems analysis of the interactions of thermal, lubricant and combustion processes of Diesel engines*, Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering, vol. 224, pp. 681-704, 2010.

[15] **A. Maiboom, X. Tauzia, J.-F. Hetet, M. Cormerais, M. Tounsi, T. Jaine and S. Blanchin**, *Various effects of EGR on combustion and emissions on an automotive DI Diesel engine: Numerical and experimental study*, SAE Paper Number 2007-01-1834, 2007.

[16] **N. Ladommatos, S. M. Abdelhalim, H. Zhao and Z. Hu**, *Effects of EGR on Heat Release in Diesel Combustion*, SAE Paper Number 980184, 1998.

[17] **D. Tomazic and A. Pfeifer**, *Cooled EGR - A must of an option for 2002/04*, SAE Paper Number 2002-01-0962, 2001.

[18] **M. Piddock, S. Akehurst, J. G. Hawley and I. Pegg**, *Engine cooling systems - Developing sound approaches for assessing system hardware changes*, in VTMS 8 - Vehicle Thermal Management Systems Conference and Exhibition, Witney, Oxford, OX28 4BN, United Kingdom, 2007, pp. 311-321.

[19] **K. Robinson**, *IC Engine coolant heat transfer studies,* (PhD Thesis), Department of Mechanical Engineering, University of Bath, 2001.

[20] **C. J. Brace, H. Burnham-Slipper, R. S. Wijetunge, N. D. Vaughan, K. Wright and D. Blight**, *Integrated Cooling systems for Passenger vehicles*, SAE Paper Number 2001-01-1248, 2001.

[21] **J. R. Wagner, M. C. Ghone, W. Dawson and E. E. Marotta**, *Coolant flow control strategies for automotive thermal management systems*, SAE Paper Number 2002-01-0713, 2002.

[22] **R. D. J. Chalgren**, *Thermal comfort and Engine warm up Optimization of a low flow advanced Thermal management system*, SAE Paper Number 2001-01-0047, 2004.

[23] **C. J. Brace, G. Hawley, S. Akehurst, M. Piddock and I. Pegg**, *Cooling system improvements - Assessing the effects on emissions and fuel economy*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 222, pp. 579-591, 2008.

[24] **D. J. Allen and M. P. Lasecki**, *Thermal Management Evolution and Controlled Coolant Flow*, SAE Paper Number 2001-01-1732, 2001.

[25] **H. H. Pang**, *Engine Thermal Management in Light Duty Diesel Engine,* (PhD Thesis), Department of Mechanical Engineering, University of Bath, Bath, 2006.

[26] **M. Chanfreau, A. Joseph, D. Butler and R. Swiatek**, *Advanced Engine Cooling Thermal Management System on a Dual Voltage 42V-14V Minivan*, SAE Paper Number 2001-01-1742, 2001.

[27] **H. Couetouse and D. Gentile**, *Cooling System control in Automotive Engines*, SAE Paper Number 920788, 1992.

[28] **H. H. Pang and C. J. Brace**, *Review of engine cooling technologies for modern engines*, Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering, vol. 218, pp. 1209-1215, Nov 2004.

[29] **I. C. Finlay, D. Harris, D. J. Boam and B. I. Parks**, *Factors Influencing Combustion Chamber Wall Temperatures in a Liquid-Cooled, Automotive, Spark Ignition Engine*, Proceedings of the Institution of Mechanical Engineers. Part D, Transport engineering, vol. 199, pp. 207-214, 1985.

[30] **M. J. Clough**, *Precision Cooling of a Four Valve per Cylinder Engine*, SAE Paper Number 931123, 1993.

[31] **A. Mulemane and R. Soman**, *CFD Based Complete Engine Cooling Jacket Development and Analysis*, SAE Paper Number 2007-01-4129, 2007.

[32] **L. You-chang, G. Xiao-hong, and C. Dan**, *Research on cooling system for 4-cylinder diesel engine*, SAE Paper Number 2007-01-2064, 2007.

[33] **L. D. I. Matteo, F. Fortunato, P. Oliva, M. Ercole, F. Mariniello, A. Montieri, S. Scala, A. Loddo and M. Martinetto**, *New and advanced cooling system concept for small cars*, in Vehicle Thermal Management 6 (VTMS 6), Brighton, United kingdom, 2003, pp. 685-694.

[34] **A. Choukroun and M. Chanfreau**, *Automatic Control of Electronic Actuators for an Optimized Engine Cooling Thermal Management*, SAE Paper Number 2001-01-1758, 2001.

[35] **M. Andre**, *Experimental study on the actual uses of the cars (E.U.R.E.V.)*, SAE Paper Number 890874, 1989.

[36] **K. Kunze, S. Wolff, I. Lade and J. Tonhauser**, *A systematic analysis of CO2 reduction by an optimized heat supply during vehicle warm up*, SAE Paper Number 2006-01-1450, 2006.

[37] **N. Ap and M. Tarquis**, *Innovative Engine cooling systems comparison*, SAE Paper Number 2005-01-1378, 2005.

[38] **W. Krause and K. H. Spies**, *Dynamic Control of the coolant temperature for reduction of fuel consumption and hydrocarbon emission*, SAE paper number 960271, 1996.

[39] **L. Jarrier, J. C. Champoussin, R. Yu and D. Gentile**, *Warm-up of a D.I. Diesel Engine: Experiment and Modelling*, SAE Paper Number 2000-01-0299, 2000.

[40] **P. J. Shayler, S. J. Christian and T. Ma**, *A Model for the investigation of Temperature, Heat flow and Friction Characteristics During Engine warm up*, SAE Paper Number 931153, 1993.

[41] **J. Moffat, M. Farid, M. Davies and A. Dris**, *The use of co-simulation to predict vehicle and engine warm-up temperature transients*, in Vehicle Thermal Management 6 (VTMS 6), Brighton, United kingdom, 2003, pp. 763-775.

[42] **J.-P. Zammit, P. J. Shayler and I. Pegg**, *Thermal coupling and eneregy flows between coolant, engine structure and lubricating oil during engine warm up*, presented at the VTMS 11 conference, Gaydon, UK, 2011.

[43] **A. J. Torregrosa, A. Broatch, P. Olmeda and C. Romero**, *Assessment of the influence of different cooling system configurations on engine warm-up, emissions and fuel consumption*, International Journal of Automotive Technology, vol. 9, pp. 447-458, Aug 2008.

[44] **P. J. Shayler, W. S. Baylis, J. P. Chick and P. Bell**, *Routes to improving heater and engine performance during warm-up*, in VTMS 4 – Vehicle Thermal Management Systems, London, UK, 1999, pp. 735-753.

[45] **D. Kay, M. Davies, J. Caine, R. McCabe and J. Theis**, *Application of a gas-to-coolant exhaust gas heat exchanger: Fuel economy and emissions benefits*, in VTMS 8 - Vehicle Thermal Management Systems Conference and Exhibition, Nottingham, United Kingdom, 2007, pp. 379-388.

[46] **G. E. Andrews, A. M. Ounzain, L. H., M. Bell, J. Tate and K. Ropkins**, *The use of a water/Lube oil heat exchanger and enhanced cooling water heating to increase water and lube oil heating rates in passenger cars for reduced fuel consumption and CO2 emissions during cold start*, SAE Paper Number 2007-01-2067, 2007.

[47] **H. De Ciutiis, T. Burgin and L. Gorlato**, *Effect of several engine encapsulation concepts on emissions, consumption and on thermal safety of a vehicle*, in VTMS 8 - Vehicle Thermal Management Systems Conference and Exhibition, Nottingham, United kingdom, 2007, pp. 389-399.

[48] **P. Revereault, C. Rouaud and A. MarchI**, *Fuel Economy and Cabin Heating Improvements Thanks to Thermal Management Solutions Installed in a Diesel Hybrid Electric Vehicle*, SAE paper number 2010-01-0800, presented at the SAE 2010 World Congress, Detroit, Michigan, 2010.

[49] **M. Jakobi, P. Hofmann and B. Geringer**, *New Heat storage Technologies for the application in future vehicles*, presented at the 32. Internationales Wiener Motorsymposium (32nd Vienna Symposium), Vienna, 2011.

[50] **R. Richter, M. Braun, D. Gosslau, R. Binnenbruck and P. Steinberg**, *Reduction of fuel consumption by predictive thermal management*, presented at the 2006 FISITA World Automotive Congress, 2006.

[51] **B. J. Hamrock, S. R. Schmid and B. O. Jacobson**, *Fundamentals of Film Fluid Lubrication*. New York: Marcel Dekker Inc., 2004.

[52] **R. Stone**, *Introduction to Internal Combustion Engines*, 3rd ed. Basingstoke: Macmillan Press LTD, 1999.

[53] **M. T. Noorman, D. N. Assanis, D. J. Patterson, S. C. Tung and S. I. Tseregounis**, *Overview of techniques for measuring friction using bench tests and fired engines*, SAE Paper Number 2000-01-1780, 2000.

[54] **Y. Wakuri, M. Soejima, Y. Ejima, T. Hamatake and T. Kitahara**, *Studies on Friction Characteristics of reciprocating Engines*, SAE Paper Number 952471, 1995.

[55] **C. C. Daniels and M. Braun**, *The friction behaviour of individual components of a spark-ignition engine during warm-up*, Tribology and Lubrication Technology, vol. 62, pp. 62-69, 2006.

[56] **D. K. W. Leong, P. J. Shayler, I. G. Pegg and M. Murphy**, *Characterizing the effect of viscosity on friction in the piston assembly of internal combustion engines*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 221, pp. 469-478, 2007.

[57] **P. J. Shayler, D. K. W. Leong and M. Murphy**, *Contributions to engine friction during cold, low speed running and the dependence on oil viscosity*, SAE Paper Number 2005-01-1654, 2005.

[58] **Y. Shimada, S. Abou, K. Okita and M. Chuubachi**, *Development of Friction Prediction Procedure and Friction Reduction Technologies for New Nissan HR and MR Engines*, SAE paper Number 2006-01-0618, 2006.

[59] **C. D. Rakopoulos, D. T. Hountalas, A. P. Koutroubousis and T. C. Zannis**, *Application and Evaluation of a Detailed Friction Model on a DI Diesel Engine with Extremely High Peak Combustion Pressures*, SAE paper Number 2001-01-0068, 2002.

[60] **R. I. Taylor and R. C. Coy**, *Improved fuel economy by lubricant design: a review*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 214, pp. 1-15, 1999.

[61] **R. I. Taylor**, *Engine friction: the influence of lubricant rheology*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 211, pp. 235-246, 1997.

[62] **Y. H. Zweiri, J. F. Whidborne, and L. D. Seneviratne**, *Instantaneous friction components model for transient engine operation*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 214, pp. 809-824, 2000.

[63] **E. Abu-Nada, I. Al-Hinti, A. Al-Sarkhi and B. Akash**, *Effect of piston friction on the performance of SI engine: A new thermodynamic approach*, Journal of Engineering for Gas Turbines and Power, vol. 130, p. 022802, 2008.

[64] **G. Livanos and N. P. Kyrtatos**, *A model of the friction losses in Diesel Engines*, SAE Paper Number 2006-01-0888, 2006.

[65] **K. J. Patton, R. G. Nitschke and J. B. Heywood**, *Development and Evaluation of a Friction Model for Spark-Ignition Engines*, SAE Paper Number 890836, 1989.

[66] **S. C. Tung and M. L. McMillan**, *Automotive tribology overview of current advances and challenges for the future*, Tribology International, vol. 37, pp. 517-536, 2004.

[67] **J. Lee, D. J. Patterson, K. M. Morrison and G. B. Schwartz**, *Friction Measurement in valve train with a roller follower*, SAE Paper Number 940589, 1994.

[68] **M. Teodorescu, D. Taraza, N. A. Henein and W. Bryzik**, *Experimental analysis of dynamics and friction in valve train systems*, SAE Paper Number 2002-01-0484, 2002.

[69] **A. Caines and R. Haycock**, *Automotive Lubricants Reference Book*, 2nd ed. Stuttgart: Robert Bosch GmbH, 1996.

[70] **U. Adler, H. Bauer, W. Bazlen and F. Dinkler**, Eds., *Automotive Handbook*. Stuttgart: Robert Bosch GmbH, 1986

[71] **G. Lechner, A. Knafl, D. Assanis, S. I. Tseregounis, M. L. McMillan, S. C. Tung, P. A. Mulawa, E. Bardasz and S. Cowling**, *Engine oil effects on the friction and emissions of a light duty, 2.2L Direct -injection- Diesel engine Part 1: Engine test results*, SAE paper Number 2002-01-2681, 2002.

[72] **S. C. Tung, M. L. McMillan, H. Gao and E. Bardasz**, *Engine Oil Effects on Friction and Wear Using 2.2L Direct Injection Diesel Engine Components for Bench Testing Part 2- Tribology Bench Test Results and Surface Analyses*, SAE Paper Number 2004-01-2005, 2004.

[73] **J. B. Heywood**, *Internal combustion engines fundamentals*: McGraw-Hill Book Company, 1988.

[74] **M. Rundo and R. Squarcini**, *Experimental Procedure for Measuring the Energy Consumption of IC Engine Lubricating Pumps during a NEDC Driving Cycle*, SAE paper number 2009-01-1919, 2009.

[75] **P. G. Evans and K. Johanson**, *The System Performance Benefits of Lubrication Flow Control*, SAE paper number 2004-01-2687, 2004.

[76] **F. Toyoda, Y. Kobayashi, Y. Miura and Y. Koga**, *Development of Variable Discharge OilPump*, SAE paper number 2008-01-0087, 2008.

[77] **D. Staley, B. Pryor and K. Gilgenback**, *Adaptation of a Variable Displacement Vane Pump to Engine Lube Oil Applications*, SAE paper number 2007-01-1567, 2007.

[78] **H. Neukirchner, M. Kramer and T. Ohnesorge**, *The controlled vane-type oilpump for oil supply on demand for passenger car engines*, SAE paper number 2002-01-1319, 2002.

[79] **C. J. Brace, J. G. Hawley, A. Cox, I. G. Pegg and R. Stark**, *The effect of variable flow oil pumps on vehicle fuel economy*, in Low-Carbon Vehicles 2009, May 20, 2009 - May 21, 2009, Westminster, London, United kingdom, 2009, pp. 219-226.

[80] **A. K. Agarwal and M. B. Varghese**, *Numerical investigations of piston cooling using oil jet in heavy duty diesel engines*, International Journal of Engine Research, vol. 7, pp. 411-421, 2006.

[81] **M. B. Varghese, S. K. Goyal and A. K. Agarwal**, *Numerical and Experimental Investigation of Oil-Jet-Cooler Piston*, SAE paper number 2005-01-1382, 2005.

[82] **P. J. Shayler, A. J. Allen, D. K. W. Leong, I. Pegg, A. J. Brown and J.-C. Cumenil**, *Characterising Lubricant Oil Viscosity to Describe Effects on Engine Friction*, SAE Paper Number 2007-01-1984, 2007.

[83] **R. I. Taylor, M. A. Brown, D. M. Thompson and J. C. Bell**, *The influence of Lubricant Rheology on Friction in the Piston ring Pack*, SAE Paper Number 941981, 1994.

[84] **T. Law, P. J. Shayler and I. Pegg**, *Investigations of sump design to improve the thermal management of oil temperature during engine warm up*, in VTMS 8 - Vehicle Thermal Management Systems Conference and Exhibition, Nottingham, United kingdom, 2007, pp. 299-309.

[85] **M. P. B. Musculus**, *On the correlation between NOx Emissions and the premix burn*, SAE Paper Number 2004-01-1401, 2004.

[86] **J. E. Dec and R. E. Canaan**, *PLIF Imaging of NO formation in a DI Diesel Engine*, SAE Paper Number 980147, 1998.

[87] **P. F. Flynn, G. L. Hunter, R. P. Durrett, L. A. Farrell and W. C. Akinyemi**, *Minimum Engine Flame Temperature Impacts on Diesel and Spark-Ignition Engine NOx Production*, SAE paper Number 2000-01-1177, 2000.

[88] **D. R. Tree and K. I. Svensson**, *Soot processes in compression ignition engines*, Progress in Energy and Combustion Science 33, pp. 272-309, 2007.

[89] **A. J. Torregrosa, P. Olmeda, J. Martin and B. Degraeuwe**, *Experiments on the influence of inlet charge and coolant temperature on performance and emissions of a DI Diesel engine*, Experimental Thermal and Fluid Science, vol. 30, pp. 633-641, 2006.

[90] **Y. Zhu, H. Zhao and N. Ladommatos**, *Computational Study of the effects of Injection Timing, EGR and Swirl Ratio on a HSDI Multi-Injection Diesel Engine Emission and Performance*, SAE Paper Number 2003-01-0346, 2003.

[91] **R. Rotondi, O. Soriano Palao and J. Helie**, *Some technical basisfor Euro6/Euro7 Diesel spray*, presented at the Diesel engines, Facing the competitiveness challenges, Rouen, France, 2010.

[92] **N. Ladommatos, S. M. Abdelhalim, H. Zhao and Z. Hu**, *The Dilution, Chemical and Thermal Effects of Exhaust Gas Recirculation on Diesel Engine Emissions - Part 4: Effects of Carbon Dioxide and Water Vapour*, SAE Paper Number 971660, 1997.

[93] **H. H. Pang, C. J. Brace and S. Akehurst**, *Potential of a Controllable Engine cooling system to reduce NOx emissions in Diesel engines*, SAE Paper Number 2004-01-0054, 2004.

[94] **W. A. Abdelghaffar, M. N. Saeed, M. M. Osman and A. I. Abdelfattah**, *Effects of coolant temperature on the performance and emissions of a diesel engine*, ASME paper number 2002-ICE-464, presented at the ASME ICE Spring technical conference,New York, NY 10016-5990, United States, 2002, pp. 187-197.

[95] **C. D. Rakopoulos, E. G. Giakoumis and D. C. Rakopoulos**, *Cylinder wall temperature effects on the transient performance of a turbocharged Diesel engine*, Energy Conversion and Management, vol. 45, pp. 2627-2638, 2004.

[96] **C. J. Brace, R. D. Burke and J. Moffa**, *Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 223, pp. 1163-1177, 2009.

[97] **R. D. Burke, C. J. Brace, and J. G. Hawley**, *Critical analysis of on-engine fuel consumption measurement*, Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering, vol. 255, pp. 829-844, 2011.

[98] **C. J. Brace, R. D. Burke and J. Moffa**, *A Design of Experiments Approach to the Control of Chassis Dynamometer Testing Error*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[99] *Engine Power Test Code~Spark Ignition and Compression ignition~Net Power Rating*, SAE Internation standard number J1349 MAR2008, 2008.

[100] **A. Lewis, C. J. Brace and A. Cox**, *The effect of forced cool down on cold start test repeatability*, SAE paper number 2009-01-1976, 2009.

[101] **K. Ropke, A. Nessler, C. Haukap, W. Baumann, B.-U. Kohler and S. Schaum**, *Model-based Methods for Engine Calibration - Quo Vadis*, presented at the 3. Internationales Symposium fur Entwicklungsmethodik. Herausforderungen im Spannungsfeld neuer Antriebskonzepte, Kurhaus Wisebaden, Germany, 2009.

[102] **M. Nebel, M.-S. Vogels, T. Combe, T. Winsel, H. Pfluegl and C. Hametner**, *Global Dynamic Models for XiL-based Calibration*, SAE paper number 2010-01-0329, 2010.

[103] **K. Shimojo, Y. Kitamura, M. Sato, M. Reumuller, R. Rojnik, M.-S. Vogels and T. Combe**, *Global dynamic modelling: a consistent approach for both diesel and gasoline engines*, presented at the 6th Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2011.

[104] **R. Dauphin, J. Chauvin and M. Castagne**, *DoEs on automated single cylinder GDI engine: towards deeper and quicker characterization of new combustion systems*, presented at the 6th Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2011.

[105] **R. Diener, T. Huber, V. Imhof, E. Kloppenburg, T. Kruse, T. Lang, U. Schulmeister and H. Ulmer**, *Challenges during the broad implementation of model-based methods in development and calibration of ECUs*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[106] **D. Sampson, I. Noell and L. Sheridan**, *Use of model based calibration in the test cell*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[107] **A. Gullitti, D. Nutter and M. Francheschi**, *Demonstration of the DoE process with software tools*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[108] **T. Dvorak, L. Malone and R. Hoekstra**, *Statistical Process control and design of experiment process improvements for the powertrain laboratory*, SAE Paper Number 2003-01-3208, 2003.

[109] **K. Roepke, A. Rosenek, M. Fischer and P. E. P. Pep**, *Practical application of DoE methods in the development of production internal combustion engines*, in International Conference on Statistics and Analytical Methods in Automotive Engineering, London, England, 2002, pp. 105-115.

[110] **M. Reger, R. Diener, V. Imhof, T. Lang, A. Powrosnik, H. Schmidt and H. Ulmer**, *Extention of classical DoE method to dynamic conditions by means of characteristic numbers*, presented at the 6th Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2011.

[111] **A. Emtage, D. Sampson and K. Ropke**, *Proposal for a Generic Interface between Test Bench Automation System and DoE Modelling Application*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[112] **G. E. P. Box, J. S. Hunter and W. G. Hunter**, *Statistics for Experimenters. Design, Innovation and Discovery*. Hoboken, New Jersey: John Wiley and sons, 2005.

[113] **A. C. Atkinson and A. N. Donev**, *Optimum Experimental Designs*. New York: Oxford University Press, 1992.

[114] **N. R. Draper and H. Smith**, *Applied Regression Analysis*, 2nd ed. New York: John Wiley and Sons Inc., 1981.

[115] **N. R. Draper and H. Smith**, *Applied Regression Analysis*, 3rd ed. New York: John Wiley and Sons Inc., 1998.

[116] **L. Eriksson, E. Johansson, N. Kettaneh-Wold, C. Wilkström and S. Wold**, *Design of Experiments: Principles and Applications*: Umetrics Academy, 2000.

[117] **K. Ropke, R. Gaitzsch, C. Haukap, M. Knaak, C. Knobel, A. Nessler, S. Schaum, U. Schoop and S. Tahl**, *DoE - Design of Experiments. Methods and applications in engine development*. Munich, Germany: Verlag Moderne Industrie, 2005.

[118] **C. Guhmann and J.-M. Riedel**, *Comparaison of Identification Methods for Nonlinear Dynamic Systems*, presented at the 6th Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2011.

[119] **W. Baumann, K. Klug, B.-U. Kohler and K. Ropke**, *Modelling of Transient Diesel Engine Emissions*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[120] **W. Baumann, T. Dreher, C. Haukap, T. Riesburg and K. Ropke**, *Nonlinear Dynamic modelling of Diesel Engines*, presented at the SIA Diesel Engine international Conference, Rouen, France, 2008.

[121] **M. Deflorian and F. Klopper**, *Design of Dynamic Experiments*, presented at the Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2009.

[122] **M. Ezzeddinne, E. Castro and R. Lengelle**, *Dynamic Design of Experiments for Engine Pollutants Emissions Modelling and Optimization*, SAE paper number 2008-01-2454, presented at the SAE Powertrains, Fuels and Lubricants Meeting, Rosemont, Illinois, 2008.

[123] **W. Baumann, S. Schaum, K. Roepke and M. Knaak**, *Excitation Signals for Nonlinear Dynamic Modelling of Combustion Engines*, presented at the Proceedings of the 17th World Congress The international Federation of Automatic control, Seoul, Korea, 2008.

[124] **M. Plint and A. Martyr**, *Engine Testing Theory and Practice*, 2nd ed. Warrendale: SAE International, 1999.

[125] **C. D. Bannister**, *Vehicle Emissions Measurement* (PhD Thesis), Department of Mechanical Engineering, University of Bath, Bath, 2007.

[126] **C. D. Bannister, J. G. Hawley, C. J. Brace, A. Cox, D. Ketcher and R. Stark**, *Further Investigations on Time Alignment*, SAE Paper Number 2004-01-1441, 2004.

[127] BS ISO 8178-1:2006. *Reciprocating internal combustion engines. Exhaust emission measurement. Test-bed measurement of gaseous and particulate exhaust emissions*, BSI.

[128] **N. Guerrassi and P. Dupraz**, *A common rail injection system for high-speed, direct-injection diesel engines*, SAE paper number 980803, presented at the SAE International Congress and Exposition, Detroit, Michigan, 1998.

[129] **M. Dzida and P. Prusakiewicz**, *The effect of temperature and pressure on the physicochemical properties of petroleum diesel oil and biodiesel fuel*, Fuel, vol. 87, pp. 1941-1948, 2008.

[130] **L. M. Rodriguez-Anton, J. Casanova-Kindelan and G. Tardajos**, *High Pressure Physical Properties of Fluids used in Diesel Injection Systems*, SAE paper number 2000-01-2046, presented at the International Spring Fuels and Lubricants meeting and exposition, Paris, 2000.

[131] **L. G. Dodge, T. J. Callahan and T. W. Ryan**, *Humidity and Temperature Correction Factors for NOx emissions from Diesel Engines*, SwRI project No. 03.30.10.06599, Southwest Research Institute, San Antonia, USA2003.

[132] BS EN ISO 8178-1:1996, *Reciprocating internal combustion engines. Exhaust emission measurement. Test-bed measurement of gaseous and particulate exhaust emissions,* BSI.

[133] **J. W. Gingrich, T. J. Callahan and L. G. Dodge**, *Humidity and Temperature Correction Factors for NOx emissions from Spark Ignited Engines*, SwRI project no. 03.10038, Southwest Research Institute, San Antonia, USA 2003.

[134] **G. Rogers and Y. Mayhew**, *Engineering Thermodynamics, Work and heat transfer*, 4th Edition ed. Harlow, UK: Addison Wesley Longman Ltd, 1992.

[135] **A. Maiboom**, *Etude Experimentale et modelisation phenomenologique de l'influence des characteristiques thermodynamiques et de la compression des gaz d'admission sur la combustion et les emissions d'un moteur Diesel automobile* (PhD Thesis), Ecole Centrale Nantes, 2007.

[136] **J. R. Sodre and S. M. C. Soares**, *Comparison of engine power correction factors for varying atmospheric conditions*, Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 25, pp. 279-284, 2003.

[137] **B. Kegl and S. Pehan**, *Reduction of Diesel Engine Emissions by Water Injection*, SAE paper Number 2001-01-3259, 2001.

[138] **D. S. Moore, G. P. McCabe and B. A. Craig**, *Introduction to the practice of statistics*. New York, USA: W. H. Freeman and company, 2009.

[139] **K. Lemon, T. Dmuchowski and B. Emaus**, *Introduction to CAN Calibration Protocol*,SAE paper Number 2000-0-0389, 2000.

[140] **A. Lewis, C. J. Brace, S. Akehurst and I. Pegg**, *A sub-miniature heat-flux sensor for use in I.C.engines*, presented at the VTMS 11, Gaydon, UK, 2011.

[141] **C. Finol**, *Heat Transfer Investigations in a Modern Diesel Engine* (PhD Thesis), Department of Mechanical Engineering, University of Bath, Bath, 2008.

[142] **E. Pipitone, A. Beccari and S. Beccari**, *Reliable TDC Position Determination: a Comparison of Different Thermodynamic Methods Through Experimental Data and Simulations*, SAE paper number 2008-36-0059, presented at the SAE Brasil 2008 Congress and Exhibit, Sao Paulo, Brazil, 2008.

[143] **R. H. Kuratle and B. Marki**, *Influencing parameters and error sources during indication on internal combustion engines*, SAE paper number 920233, 1992.

[144] **A. L. Randolph**, *Methods of Processing Cylinder-Pressure Transducer Signals to Maximize Data Accuracy*, SAE Paper 900170, 1990.

[145] **G. F. Hohenberg**, *Advanced Approaches for Heat Transfer Calculations*, SAE Paper No. 790825, 1979.

[146] **F. Hofmann**, *Fundemental Principles of Electromagnetic Flow Measurement*, 3rd Edition, Krohne, Duisburg, 2003, available from: http://www.krohne.com/fileadmin/media-lounge/PDF-Download/Magmeters/HB_MID_e.pdf (accessed June 2010).

[147] **S. Son and A. E. Kolasa**, *Estimating Actual Exhaust Gas Temperature from Raw Thermocouple Measurements Acquired During Transient and Steady State Engine Dynamometer Tests*, SAE paper Number 2007-01-0335, 2007.

[148] **R. D. Burke and C. J. Brace**, *The Effects of Engine Thermal Conditions on Performance, Emissions and Fuel Consumption*, SAE Paper Number 2010-01-0802, presented at the SAE World Congress 2010, Detroit, USA, 2010.

[149] **R. D. Burke, C. J. Brace, A. Cox, A. Lewis, J. G. Hawley, I. Pegg and R. Stark**, *Systems approach to the improvement of engine warm-up behaviour*, Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering, vol. 225(2), pp. 190-205, 2011.

[150] **E. Pipitone**, *A New Simple Friction Model for S. I. Engine*, SAE paper Number 2009-01-1984, 2009.

[151] **J. G. Hawley, C. D. Bannister, C. J. Brace, I. Pegg, J. C. Dumenil and A. Brown**, *Developing robust process control for determining the effect of lubricating oil properties and additives on diesel engine fuel economy*, presented at the Additives 2007: Application for Future Transport, RSC International Meeting, London, 2007.

[152] **T. M. Morton and S. Knott**, *Radial Basis functions for engine modelling*, International Conference on Statistics and Analytical Methods in Automotive Engineering, London, England, 2002, pp. 43-51.

[153] **A. Le Negrate, M. El Fassi and G. Cavallaro**, *Local oil flow measurement in Diesel engine*, presented at the Diesel engines, Facing the competitiveness challenges, Rouen, France, 2010.

[154] **F. P. Incropera and D. P. De Witt**, *Introduction to Heat transfer*, 2nd ed. New York: John Wiley and sons Inc., 1985.

[155] **R. Burke, C. J. Brace, S. Akehurst, B. Fath, W. Baumann and R. Waschecck**, *Practical approach to the thermo-dynamic modelling of Diesel engine*, presented at the 6th Design of Experiments (DoE) in Engine Development, Berlin, Germany, 2011.

[156] The Mathworks. (2011, 07/07/2011). *Matlab product documentation, Signal processing toolbox: Chirp - Swept frequency cosine*. Available from: http://www.mathworks.com/help /toolbox/signal/chirp.html (accessed July 2011)

[157] Qhull (1995, 01/08/2011). *Qhull*. Available: http://www.qhull.org (accessed August 2011)